



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 99576

TO: Stephen Rawlings
Location: CM1/8E17/8E12
Art Unit: 1642
Tuesday, August 12, 2003
Case Serial Number: 09/492764

From: Paul Schulwitz
Location: Biotech-Chem Library
CM1-6B06
Phone: 305-1954

paul.schulwitz@uspto.gov

Search Notes

Examiner Rawlings,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(703)305-1954

*Will
Amend
gen electron?*

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:55:43 ; Search time 83 Seconds
(without alignments)
11.474 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLKTK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_15Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	26	96.3	54	22 ABG5252
2	26	96.3	54	22 ABB40087
3	26	96.3	54	22 AAM60843
4	26	96.3	54	22 AAM73525
5	26	96.3	54	22 AAM33721
6	26	96.3	54	23 ABG43386
7	26	96.3	93	22 AAO05971
8	26	96.3	113	21 AAG07848
9	26	96.3	114	21 AAG07847

10	26	96.3	114	21 AAG09625
11	26	96.3	114	21 AAG48379
12	26	96.3	130	22 ABG23720
13	26	96.3	138	21 AAG09624
14	26	96.3	138	21 AAG48378
15	26	96.3	169	21 AAG09623
16	26	96.3	169	21 AAG48377
17	26	96.3	174	23 ABJ11303
18	26	96.3	174	23 ABG77401
19	26	96.3	266	22 ABB50188
20	26	96.3	266	22 AAM93254
21	26	96.3	293	20 AAW88324
22	26	96.3	350	23 ABP64568
23	26	96.3	351	23 ABB97313
24	26	96.3	355	23 ABP42011
25	26	96.3	480	23 ABB53262
26	26	96.3	543	23 ABG65125
27	26	96.3	577	23 ABU65124
28	26	96.3	577	24 ABB82677
29	26	96.3	583	23 ABB53263
30	26	96.3	718	16 AAR85865
31	26	96.3	719	15 AAR81481
32	26	96.3	719	15 AAR522955
33	26	96.3	719	22 ABB63831
34	26	96.3	719	22 ABB66968
35	26	96.3	774	22 ABB63440
36	26	96.3	896	19 AAW41252
37	25	92.6	6	21 AAB12856
38	25	92.6	6	21 AAB12865
39	25	92.6	9	23 AAU10434
40	25	92.6	9	23 AAU10454
41	25	92.6	12	17 AAR98426
42	25	92.6	14	21 AAB12847
43	25	92.6	15	22 AAB10407
44	25	92.6	50	22 AAM88521
45	25	92.6	52	21 AAG01675

ALIGNMENTS

RESULT 1
ABG525252
ID ABG525252 standard; Peptide; 54 AA.
XX AC ABG525252;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver peptide, SEQ ID No 33900.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US006664.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 54 AA;
 SQ Query Match 96.3%; Score 26; DB 22; Length 54;
 Best Local Similarity 83.3%; Pred. No. 35;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 | | | | |
 Db 34 PSLKTK 39

RESULT 3

AAAM60843
 ID AAAM60843 standard; Protein; 54 AA.
 XX
 AC AAAM60843;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32948.

XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 XN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -
 XX Example 4; SEQ ID NO: 32948; 650pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX brain. They can be used to measure gene expression in brain cell samples,
 XX which may enable the diagnosis and improved treatment of nervous system
 XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 XX epilepsy and cancers. The present sequence is a protein encoded by one of
 XX the probes of the invention.

XX Sequence 54 AA;
 SQ Query Match 96.3%; Score 26; DB 22; Length 54;

XX WPI; 2001-488998/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -
 XX Claim 27; SEQ ID No 33900; 658pp; English.
 XX
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/ fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult
 XX liver. (I) may be used for predicting, measuring and displaying gene
 XX expression in samples derived from human adult liver. The genes
 XX identified may be involved in genetic liver diseases such as cirrhosis,
 XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 XX is associated with coronary heart disease. ABG47348-ABG59930 represent
 XX human liver single exon encoded peptides of the invention.
 XX Note: The sequence information for this patent does not appear in the
 XX printed specification but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 54 AA;
 SQ Query Match 96.3%; Score 26; DB 22; Length 54;
 Best Local Similarity 83.3%; Pred. No. 35;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 | | | | |
 Db 34 PSLKTK 39

RESULT 2

ABB40087
 ID ABB40087 standard; Peptide; 54 AA.
 XX
 AC ABB40087;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #7593 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX Homo sapiens.
 XX
 XN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human fetal liver -
 XX Claim 27; SEQ ID NO 32722; 639pp + sequence listing; English.

XX Sequence 54 AA;
 SQ Query Match 96.3%; Score 26; DB 22; Length 54;

Best Local Similarity 83.3%; Pred. No. 35; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
| | | | |
Db 34 PSLKTK 39

RESULT 4
AAM73525
ID AAM73525 standard; Protein; 54 AA.

XX AC AAM73525;
XX 06-NOV-2001 (first entry)
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33831.
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX Homo sapiens.

XX WO200157276-A2.
XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX Example 4; SEQ ID NO: 33831; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.

XX Sequence 54 AA;

Query Match 96.3%; Score 26; DB 22; Length 54;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
| | | | |
Db 34 PSLKTK 39

RESULT 5
AAM33721
ID AAM33721 standard; Protein; 54 AA.

XX AAM33721;

XX 17-OCT-2001 (first entry)

XX Peptide #7758 encoded by probe for measuring placental gene expression.
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.

XX Homo sapiens.
XX WO200157272-A2.
XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488997/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX Claim 27; SEQ ID No 33990; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;
XX see AA131315-AA15746). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.

XX Sequence 54 AA;

Query Match 96.3%; Score 26; DB 22; Length 54;
Best Local Similarity 83.3%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
| | | | |
Db 34 PSLKTK 39

RESULT 6
ABG43386
ID ABG43386 standard; Peptide; 54 AA.

XX ABG43386;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 33051.
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangiomyomatosis; Karaguner syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.

XX Homo sapiens.

XX

34 PSLKTK 39

Db

RESULT 7

AAO05971 AAO05971 standard; Protein; 93 AA.

XX AC AAO05971;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 19863.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI85902.

XX PT Isolated nucleic acids and polypeptides, useful for preventing

XX PT diagnosing and treating e.g. leukaemia, inflammation and immune

XX PT disorders -

XX PS Claim 20; SEQ ID NO 19863; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 93 AA;

Query Match 96.3%; Score 26; DB 22; Length 93;

Best Local Similarity 83.3%; Pred. No. 61;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6

Db 7 PSLKTK 12

RESULT 8

AAG07848 AAG07848 standard; Protein; 113 AA.

XX ID AAG07848

XX ID AAG07848 standard; Protein; 113 AA.

XX

PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US00665.

XX PR 04-FEB-2000; 2000US-#80312P.

XX PR 26-MAY-2000; 2000US-207456P.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-234687P.

XX PR 27-SEP-2000; 2000US-236359P.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to

XX PT measure gene expression in human lung samples -

XX PS Claim 27; SEQ ID No 33051; 634pp; English.

XX CC The invention relates to a spatially-addressable set of single exon

XX CC nucleic acid probes for measuring gene expression in a sample derived

XX CC from human lung comprising single exon nucleic acid probes having one of

XX CC 12614 nucleic acid sequences mentioned in the specification, or their

XX CC complements or the 12387 open reading frames derived from the 12614

XX CC probes. Also included are a microarray comprising the novel set of

XX CC probes; the novel set of probes which hybridise at high stringency to a

XX CC nucleic acid expressed in the human lung; measuring gene expression in a

XX CC sample derived from human lung, comprising (a) contacting the array with

XX CC a collection of detectably labeled nucleic acids derived from human lung

XX CC mRNA, and (b) measuring the label detectably bound to each probe of

XX CC the array; identifying exons in a eukaryotic genome, comprising

XX CC (a) algorithmically predicting at least one exon from genomic sequences

XX CC of the eukaryote; and (b) detecting specific hybridisation of detectably

XX CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

XX CC having a fragment identical to the predicted exon, the probe is included

XX CC in the above mentioned microarray; assigning exons to a single gene.

XX CC comprising (a) identifying exons from genomic sequence by the method

XX CC above and (b) measuring the expression of each of the exons in several

XX CC tissues and/or cell types using hybridisation to a single exon

XX CC microarrays having a probe with the exon, where a common pattern of

XX CC expression of the exons in the tissues and/or cell types indicates that

XX CC the exons should be assigned to a single gene; a peptide comprising one

XX CC of 12011 sequences, mentioned in the specification, or encoded by the

XX CC genes/probes/open reading frames (ORF). The probes are used for gene

XX CC expression analysis, and for identifying exons in a gene, particularly

XX CC using human lung derived mRNA and for the study of lung diseases

XX CC such as asthma, lung cancer, chronic obstructive pulmonary disease

XX CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

XX CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

XX CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

XX CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

XX CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic

XX CC and hyaline membrane disease. The present sequence is a peptide/protein

XX CC encoded by a single exon probe of the invention.

XX CC Note: The sequence data for this patent did not form part

XX CC of the printed specification, but was obtained in electronic

XX CC format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 54 AA;

Query Match 96.3%; Score 26; DB 23; Length 54;

Best Local Similarity 83.3%; Pred. No. 35;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6

Db 7 PSLKTK 12

RESULT 8

AAG07848 AAG07848 standard; Protein; 113 AA.

XX ID AAG07848

XX ID AAG07848 standard; Protein; 113 AA.

XX

AC AAG07848; 99US-0139463.
 XX 17-OCT-2000 (first entry) 99US-0139750.
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 5157. 99US-0139763.
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 5157. 99US-0139817.
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 5157. 99US-0139899.
 KW Protein identification; signal transduction pathway; metabolic pathway; 99US-0140353.
 KW hybridisation assay; genetic mapping; gene expression control; promoter; 99US-0140354.
 KW termination sequence. 99US-0140695.
 XX Arabidopsis thaliana. 99US-0140823.
 XX Arabidopsis thaliana. 99US-0140991.
 XX Arabidopsis thaliana. 99US-0141287.
 XX Arabidopsis thaliana. 99US-0141842.
 XX Arabidopsis thaliana. 99US-0142154.
 XX Arabidopsis thaliana. 99US-0142055.
 PD Arabidopsis thaliana. 99US-0142390.
 XX Arabidopsis thaliana. 99US-0142803.
 XX Arabidopsis thaliana. 99US-0142920.
 XX Arabidopsis thaliana. 99US-0142977.
 XX Arabidopsis thaliana. 99US-0143542.
 XX Arabidopsis thaliana. 99US-0143624.
 XX Arabidopsis thaliana. 99US-0144005.
 XX Arabidopsis thaliana. 99US-0144085.
 XX Arabidopsis thaliana. 99US-0144086.
 XX Arabidopsis thaliana. 99US-0144325.
 XX Arabidopsis thaliana. 99US-0144331.
 XX Arabidopsis thaliana. 99US-0144332.
 XX Arabidopsis thaliana. 99US-0144333.
 XX Arabidopsis thaliana. 99US-0144334.
 XX Arabidopsis thaliana. 99US-0144335.
 XX Arabidopsis thaliana. 99US-0144352.
 XX Arabidopsis thaliana. 99US-0144632.
 XX Arabidopsis thaliana. 99US-0144684.
 XX Arabidopsis thaliana. 99US-0144814.
 XX Arabidopsis thaliana. 99US-0145086.
 XX Arabidopsis thaliana. 99US-0145088.
 XX Arabidopsis thaliana. 99US-0145085.
 XX Arabidopsis thaliana. 99US-0145087.
 XX Arabidopsis thaliana. 99US-0145089.
 XX Arabidopsis thaliana. 99US-0145192.
 XX Arabidopsis thaliana. 99US-0145145.
 XX Arabidopsis thaliana. 99US-0145218.
 XX Arabidopsis thaliana. 99US-0145224.
 XX Arabidopsis thaliana. 99US-0145276.
 XX Arabidopsis thaliana. 99US-0145913.
 XX Arabidopsis thaliana. 99US-0145918.
 XX Arabidopsis thaliana. 99US-0145919.
 XX Arabidopsis thaliana. 99US-0145951.
 XX Arabidopsis thaliana. 99US-0146386.
 XX Arabidopsis thaliana. 99US-0146388.
 XX Arabidopsis thaliana. 99US-0147038.
 XX Arabidopsis thaliana. 99US-0147204.
 XX Arabidopsis thaliana. 99US-0147302.
 XX Arabidopsis thaliana. 99US-0147192.
 XX Arabidopsis thaliana. 99US-0147260.
 XX Arabidopsis thaliana. 99US-0147303.
 XX Arabidopsis thaliana. 99US-0147416.
 XX Arabidopsis thaliana. 99US-0147493.
 XX Arabidopsis thaliana. 99US-0147935.
 XX Arabidopsis thaliana. 99US-0148171.
 XX Arabidopsis thaliana. 99US-0148319.
 XX Arabidopsis thaliana. 99US-0148341.
 XX Arabidopsis thaliana. 99US-0148565.
 XX Arabidopsis thaliana. 99US-0148684.
 XX Arabidopsis thaliana. 99US-0149368.
 XX Arabidopsis thaliana. 99US-0149175.
 XX Arabidopsis thaliana. 99US-0149426.
 XX Arabidopsis thaliana. 99US-0149722.
 XX Arabidopsis thaliana. 99US-0149723.
 XX Arabidopsis thaliana. 99US-0149923.
 XX Arabidopsis thaliana. 99US-0149902.
 XX Arabidopsis thaliana. 99US-0149930.
 XX Arabidopsis thaliana. 99US-0150566.

hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

Accession	Score	DB 21	Length	Indels	Mismatches	Gaps
99US-0150884	96.3%		113			
99US-0151065	83.3%					
99US-0151066						
99US-0151080						
99US-0151303						
99US-0151438						
99US-0151930						
99US-0152363						
99US-0153070						
99US-0153758						
99US-0154018						
99US-0154039						
99US-0154779						
99US-0155139						
99US-0155486						
99US-0155659						
99US-0156458						
99US-0156596						
99US-0157117						
99US-0157753						
99US-0157865						
99US-0158029						
99US-0158232						
99US-0158369						
99US-0159293						
99US-0159294						
99US-0159295						
99US-0159329						
99US-0159330						
99US-0159331						
99US-0159637						
99US-0159638						
99US-0159584						
99US-0160741						
99US-0160767						
99US-0160768						
99US-0160770						
99US-0160814						
99US-0160815						
99US-0160980						
99US-0160981						
99US-0160989						
99US-0161404						
99US-0161405						
99US-0161406						
99US-0161359						
99US-0161360						
99US-0161361						
99US-0161920						
99US-0161992						
99US-0161993						
99US-0162142						
99US-0121825						
99US-0123180						
99US-0123548						
99US-0125788						
99US-0126264						
99US-0126785						
99US-0127462						
99US-0128234						
99US-0128714						
99US-0129845						
99US-0130077						
99US-0130449						
99US-0130510						
99US-0130891						
99US-0131449						
99US-0132048						
99US-0132407						
99US-0132484						
99US-0132485						
99US-0132486						
99US-0132487						
99US-0132863						
99US-0134256						
99US-0134218						
99US-0134219						
99US-0134221						
99US-0134370						
99US-0134768						
99US-0134941						
99US-0135124						
99US-0135353						
99US-0135629						
99US-0136021						
99US-0136392						
99US-0136782						
99US-0137222						
99US-0137528						
99US-0137502						
99US-0137724						
99US-0138094						
99US-0138540						
99US-0138847						
99US-0139119						
99US-0139452						
99US-0139453						
99US-0139492						
99US-0139454						
99US-0139455						
99US-0139456						
99US-0139457						
99US-0139458						
99US-0139459						
99US-0139460						
99US-0139461						
99US-0139462						
99US-0139463						
99US-0139750						
99US-0139763						
99US-0139817						
99US-0139899						
99US-0140353						
99US-0140354						

KW Arabidopsis thaliana.
 XX EPI033405-A2.
 OS 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 XX 09-MAR-1999; 99US-0123548.
 XX 23-MAR-1999; 99US-0125788.
 XX 25-MAR-1999; 99US-0126264.
 XX 29-MAR-1999; 99US-0126785.
 XX 01-APR-1999; 99US-0127462.
 XX 06-APR-1999; 99US-0128234.
 XX 08-APR-1999; 99US-0128714.
 XX 16-APR-1999; 99US-0129845.
 XX 19-APR-1999; 99US-0130077.
 XX 21-APR-1999; 99US-0130449.
 XX 23-APR-1999; 99US-0130510.
 XX 28-APR-1999; 99US-0130891.
 XX 30-APR-1999; 99US-0131449.
 XX 04-MAY-1999; 99US-0132048.
 XX 30-APR-1999; 99US-0132407.
 XX 05-MAY-1999; 99US-0132484.
 XX 06-MAY-1999; 99US-0132485.
 XX 06-MAY-1999; 99US-0132486.
 XX 07-MAY-1999; 99US-0132487.
 XX 11-MAY-1999; 99US-0132863.
 XX 14-MAY-1999; 99US-0134256.
 XX 14-MAY-1999; 99US-0134218.
 XX 14-MAY-1999; 99US-0134219.
 XX 14-MAY-1999; 99US-0134221.
 XX 14-MAY-1999; 99US-0134370.
 XX 14-MAY-1999; 99US-0134768.
 XX 19-MAY-1999; 99US-0134941.
 XX 20-MAY-1999; 99US-0135124.
 XX 21-MAY-1999; 99US-0135353.
 XX 24-MAY-1999; 99US-0135629.
 XX 25-MAY-1999; 99US-0136021.
 XX 27-MAY-1999; 99US-0136392.
 XX 28-MAY-1999; 99US-0136782.
 XX 01-JUN-1999; 99US-0137222.
 XX 03-JUN-1999; 99US-0137528.
 XX 04-JUN-1999; 99US-0137502.
 XX 07-JUN-1999; 99US-0137724.
 XX 08-JUN-1999; 99US-0138094.
 XX 10-JUN-1999; 99US-0138540.
 XX 10-JUN-1999; 99US-0138847.
 XX 14-JUN-1999; 99US-0139119.
 XX 16-JUN-1999; 99US-0139452.
 XX 16-JUN-1999; 99US-0139453.
 XX 17-JUN-1999; 99US-0139492.
 XX 18-JUN-1999; 99US-0139454.
 XX 18-JUN-1999; 99US-0139455.
 XX 18-JUN-1999; 99US-0139456.
 XX 18-JUN-1999; 99US-0139457.
 XX 18-JUN-1999; 99US-0139458.
 XX 18-JUN-1999; 99US-0139459.
 XX 18-JUN-1999; 99US-0139460.
 XX 18-JUN-1999; 99US-0139461.
 XX 18-JUN-1999; 99US-0139462.
 XX 18-JUN-1999; 99US-0139463.
 XX 18-JUN-1999; 99US-0139750.
 XX 18-JUN-1999; 99US-0139763.
 XX 21-JUN-1999; 99US-0139817.
 XX 22-JUN-1999; 99US-0139899.
 XX 23-JUN-1999; 99US-0140353.
 XX 23-JUN-1999; 99US-0140354.

QY 1 PXLKTK 6
 Db 86 PSLKTK 91
 RESULT 9
 AAG07847
 ID AAG07847 standard; Protein; 114 AA.
 AC AAG07847;
 XX 17-OCT-2000 (first entry)
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 5156.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW Arabidopsis thaliana protein fragment SEQ ID NO: 5156.

PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161991.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.3%; Score 26; DB 21; Length 114;
 Best Local Similarity 83.3%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 87 PSLKTK 92

RESULT 10
 AAG09625
 ID AAG09625 standard; Protein; 114 AA.
 XX
 AC AAG09625;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 7630.

KW Protein identification; signal transduction pathway; metabolic pathway;
 hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX

PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 XX 09-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139617.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147939.
 PR 10-AUG-1999; 99US-0148171.
 PR 10-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0158294.
 PR 13-OCT-1999; 99US-0158295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0158637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.3%; Score 26; DB 21; Length 114;
 Best Local Similarity 83.3%; Pred.No. 76;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 86 PSLKTK 91

RESULT 11
 AAG48379
 ID AAG48379 standard; Protein; 114 AA.

XX AAG48379;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 61086.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144332.
 PR 20-JUL-1999; 99US-0144333.
 PR 20-JUL-1999; 99US-0144334.
 PR 20-JUL-1999; 99US-0144335.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 23-JUL-1999; 99US-0145276.
 PR 26-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.3%; Score 26; DB 21; Length 114;
 Best Local Similarity 83.3%; Pred. No. 76;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 86 PSLKTK 91

RESULT 12
 ABG23720
 ID ABG23720 standard; Protein; 130 AA.

XX AC ABG23720;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #23711.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.
 XX FN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 XX DR N-PSDB; AAS87907.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

PS Claim 20; SEQ ID No 54079; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AEG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 130 AA;

Query Match 96.3%; Score 26; DB 22; Length 130;

Best Local Similarity 83.3%; Pred. NO. 87;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLTKK 6

Db 74 PFLTKK 79

RESULT 13

AAG09624 ID AAG09624 standard; Protein; 138 AA.

XX AC AAG09624;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7629.

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

```

```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match Similarity 96.3%; Score 26; DB 21; Length 138;

Best Local Similarity 83.3%; Fred. No. 93;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 PXLKTK 6
        |  |  |  |
Db      110 PSLKTK 115

```

RESULT 14

```

AAG48378
ID AAG48378 standard; Protein; 138 AA.
XX AAG48378;
AC AAG48378;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61085.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.

```

PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145088.
 PR 21-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149829.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.

Tue Aug 12 09:50:57 2003

PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.3%; Score 26; DB 21; Length 138;
 Best Local Similarity 83.3%; Pred. No. 93;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 110 PSLKTK 115

RESULT 15
 AAG09623
 ID AAG09623 standard; Protein; 169 AA.
 AC AAG09623;
 XX
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 7628.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX
 XX EP1033405-A2.
 PN
 XX
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-0301439.
 XX
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143242.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145216.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161921.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.3%; Score 26; DB 21; Length 169;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 PXLKTK 6
 Db 141 PSLKTK 146

Search completed: August 4, 2003, 16:05:00
 Job time : 86 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 16:03:18 ; Search time 39 Seconds
(without alignments)
14.795 Million cell updates/sec

Title: US-09-492-764B-20

Perfect score: 27

Sequence: 1 PXLKTK 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	26	96.3	242	2	T17575	procollagen-prolin
2	26	96.3	293	2	S22613	abequose synthase
3	26	96.3	338	2	S58324	sporulation-specif
4	26	96.3	542	2	S20466	hypothetical prote
5	26	96.3	719	2	A30047	enhancer of split
6	26	96.3	764	2	T21128	hypothetical prote
7	26	96.3	849	2	D883168	protein F19H8.4 [i
8	26	96.3	899	2	T11578	probable lipoxigen
9	26	96.3	1435	1	BVBYL1	guanine nucleotide
10	25	92.6	66	2	B87379	ribosomal protein
11	25	92.6	67	2	F71666	ribosomal protein
12	25	92.6	75	2	F64423	hypothetical prote
13	25	92.6	89	2	F97816	50S ribosomal prot
14	25	92.6	142	2	C75271	hypothetical prote
15	25	92.6	160	2	T44368	hypothetical prote
16	25	92.6	191	2	S29884	Ribosomal protein
17	25	92.6	231	2	C43727	conserved hypoteth
18	25	92.6	242	2	T34384	hypothetical prote
19	25	92.6	256	2	AE2019	hypothetical prote
20	25	92.6	264	1	G69884	conserved hypoteth
21	25	92.6	269	2	S61633	YNT20 protein - ye
22	25	92.6	273	2	T37841	probable transloca
23	25	92.6	281	2	AF1359	L-alanyl-D-glutam
24	25	92.6	281	2	S69799	L-alanyl-D-glutam
25	25	92.6	312	2	D90008	hypothetical prote
26	25	92.6	327	2	A13245	conserved hypoteth
27	25	92.6	343	2	JH0442	allantoicase (EC 3
28	25	92.6	375	2	H71121	probable alcohol d
29	25	92.6	403	2	T36019	phosphoglycerate k

30	25	92.6	416	2	S65221	hypothetical prote
31	25	92.6	421	2	G71638	ampg protein (ampG
32	25	92.6	425	2	AF2246	phosphoribosylform
33	25	92.6	429	2	AB2456	glycoyltransferas
34	25	92.6	455	2	AH2058	hypothetical prote
35	25	92.6	479	2	T44326	hypothetical prote
36	25	92.6	556	2	S67097	probable membrane
37	25	92.6	567	2	AG2008	hypothetical prote
38	25	92.6	697	2	G70133	flagellar biosynth
39	25	92.6	713	2	A12514	ABC transporter AT
40	25	92.6	720	2	G70320	aldehyde dehydrog
41	25	92.6	770	2	I49508	tsGF3 p91-related
42	25	92.6	770	2	A54444	DNA-binding protei
43	25	92.6	896	2	AE1514	conserved membrane
44	25	92.6	901	2	A48653	phage infection pr
45	25	92.6	923	2	G83826	hypothetical prote

ALIGNMENTS

RESULT 1

T17575
 procollagen-proline dioxygenase alpha chain-like protein A85R - Chlorella virus PBCV-1
 C;Species: Chlorella virus PBCV-1
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T17575
 R;Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL data Library, May 1999
 A;Reference number: Z18806
 A;Accession: T17575
 A;Status: preliminary; translated from GB/EMBL/DBBJ
 A;Molecule type: DNA
 A;Residues: 1-242 <GRA>
 A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96453.1
 A;Experimental source: specific host Chlorella strain NC64A
 C;Genetics: ;
 A;Gene: A85R

Query Match 96.3%; Score 26; DB 2; Length 242;
 Best Local Similarity 83.3%; Pred. No. 51;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 Db 189 PTLKTK 194

RESULT 2

S22613
 abequose synthase - Salmonella choleraesuis (strain M67)
 C;Species: Salmonella choleraesuis
 A;Variety: Strain M67
 C;Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 17-Mar-2000
 R;Brown, P.K.; Romana, L.K.; Reeves, P.R.
 Mol. Microbiol. 6 1385-1394, 1992
 A;Title: Molecular analysis of the rfb gene cluster of Salmonella serovar muenchen (str. S22613; MUID:92349966; PMID:1379320)
 A;Reference number: S22613
 A;Accession: S22613
 A;Molecule type: DNA
 A;Residues: 1-293 <BRO>
 A;Cross-references: EMBL:X61917; NID:g47004; PIDN:CAA43918.1; PID:g47005
 A;Experimental source: strain M67 serovar muenchen
 A;Note: the authors did not translate the codon for residue 293
 C;Genetics: ;
 A;Gene: rfbJ
 C;Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology

Query Match 96.3%; Score 26; DB 2; Length 293;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

A;Cross-references: EMBL:X64799; NID:g2722; PIDN:CAA46025.1; PID:g2723

Query Match 96.3%; Score 26; DB 2; Length 542; Best Local Similarity 83.3%; Pred. No. 1.2e+02; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 113 PALKTK 118

RESULT 5
A30047
enhancer of split protein - fruit fly (Drosophila melanogaster)
N;Alternate names: neurogenic repetitive locus protein
C;Species: Drosophila melanogaster
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 26-May-2000
C;Accession: A30047
R;Hartley, D.A.; Preiss, A.; Artavanis-Tsakonas, S.
Cell 55, 785-795, 1988
A;Title: A deduced gene product from the Drosophila neurogenic locus, Enhancer of split
A;Reference number: A30047; MUID:89051868; PMID:3142687
A;Accession: A30047
A;Molecule type: mRNA
A;Residues: 1-719 <HAR>
A;Cross-references: GB:M20571; NID:gl57364; PID:gl57365
C;Genetics:

A;Gene: FlyBase:gro
A;Cross-references: FlyBase:FBgn0001139
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: nucleus; Phosphoprotein
F;475-508/Domain: WD repeat homology <WD1>
F;561-594/Domain: WD repeat homology <WD2>
F;643-676/Domain: WD repeat homology <WD3>
F;684-717/Domain: WD repeat homology <WD4>

Query Match 96.3%; Score 26; DB 2; Length 719;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 296 PSLKTK 301

RESULT 6
T21128
hypothetical protein F19H8.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21128
R;Steward, C.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19379
A;Accession: T21128
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-764 <WIL>
A;Cross-references: EMBL:Z933378; PIDN:CAR07585.2; GSPDB:GN00020; CESP:F19H8.4
A;Experimental source: clone F19H8
C;Genetics:
A;Gene: CESP:F19H8.4
A;Map position: 2
A;Introns: 30/1; 61/3; 134/2; 340/2; 393/3; 426/2; 464/3; 598/3; 646/3

Query Match 96.3%; Score 26; DB 2; Length 764;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 312 PSLKTK 317

QY 1 PXLKTK 6
| | | | |
Db 103 PALKTK 108

RESULT 3
S58324
sporulation-specific protein SPS4 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O6120; protein YOR313C
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jul-2000
C;Accession: S58324; A25391; S67219; S71993
R;Pearson, B.M.; Hernando, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
submitted to the EMBL Data Library, August 1995
A;Reference number: S58318
A;Accession: S58324
A;Molecule type: DNA
A;Residues: 1-338 <PEA>
A;Cross-references: EMBL:X90565; NID:g940836; PID:g940844
R;Garber, A.T.; Segall, J.
Mol. Cell. Biol. 6, 4478-4485, 1986

A;Title: The SPS4 gene of Saccharomyces cerevisiae encodes a major sporulation-specific
A;Reference number: A25391; MUID:87089807; PMID:3540611
A;Accession: A25391
A;Molecule type: DNA
A;Residues: 1-106; H, 108-338 <GAR>
A;Cross-references: EMBL:M14684
R;Pearson, B.M.; Hernando, Y.; Kalogeropoulos, A.; Schweizer, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67213
A;Accession: S67219
A;Molecule type: DNA
A;Residues: 1-338 <PEW>
A;Cross-references: EMBL:T75221; NID:gl420687; PID:g252145; PID:gl420688; MIPS:YOR313C
A;Experimental source: strain S288C
R;Pearson, B.M.; Hernando, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
Yeast 12, 1021-1031, 1996
A;Title: Sequencing of a 35.71 kb DNA segment on the right arm of yeast chromosome XV re
A;Reference number: S71986; MUID:97051589; PMID:8986266
A;Accession: S71993
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-338 <PEF>
A;Cross-references: EMBL:X90565; NID:g940836; PIDN:CAA62168.1; PID:g940844
A;Notice: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C;Genetics:
A;Gene: SGD:SPS4
A;Cross-references: SGD:S0005940; MIPS:YOR313C
A;Map position: 15R

Query Match 96.3%; Score 26; DB 2; Length 338;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 135 PSLKTK 140

RESULT 4
S20466
hypothetical protein - fungus (Fusarium oxysporum)
C;Species: Fusarium oxysporum
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jul-2000
C;Accession: S20466
R;Daboussi, M.J.; Langin, T.; Brygoo, Y.
Mol. Gen. Genet. 232, 12-16, 1992
A;Title: Foc1, a new family of fungal transposable elements.
A;Reference number: S20466; MUID:92204124; PMID:1313143
A;Accession: S20466
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-542 <DAB>

RESULT 7
 D88368
 Protein F19H8.4 [imported] - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C;Accession: D88368
 R;Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
 A;Reference number: A75000; MUID:99069613; PMID:9851916
 A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans
 C;Date: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A;Accession: D88368
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-849 <STO>
 A;Cross-references: GB:chr_II; PIDN:CAB07585.1; GSPDB:GN00020; CESP:F19H8
 A;Note: cDNA EST Yk238g11.5 comes from this gene
 C;Genetics:
 A;Gene: F19H8.4
 A;Map position: 2

Query Match 96.3%; Score 26; DB 2; Length 849;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | ||||
 Db 397 PSLKTK 402

RESULT 8
 T11578
 Probable lipoygenase (EC 1.13.11.12) CPRD46, drought-inducible - cowpea
 C;Species: Vigna unguiculata (cowpea)
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 20-Jun-2000
 C;Accession: T11578
 R;Iuchi, S.; Yamaguchi-Shinozaki, K.; Urao, T.; Shinozaki, K.
 J. Plant Res. 109, 415-424, 1996
 A;Title: Characterization of two cDNAs for novel drought-inducible genes in the highly drought-tolerant cowpea
 A;Reference number: Z17293
 A;Accession: T11578
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-899 <IUC>
 A;Cross-references: EMBL:D88122
 C;Function:
 A;Description: catalyzes the hydroperoxidation of specific unsaturated fatty acids
 C;Superfamily: lipoygenase
 C;Keywords: fatty acid oxidation; oxidoreductase

Query Match 96.3%; Score 26; DB 2; Length 899;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | ||||
 Db 726 PALKTK 731

RESULT 9
 BVBYL1
 Guanine nucleotide-releasing factor LTE1 - yeast (Saccharomyces cerevisiae)
 N;Alternate names: protein YAL024c
 C;Species: Saccharomyces cerevisiae
 C;Date: 31-Mar-1991 #sequence_revision 03-Nov-1995 #text_change 16-Jun-2000
 C;Accession: S51997; S45454; S43456; S05869; S46646
 R;Bussey, H.; Kaback, D.B.; Zhong, W.; Vo, D.T.; Clark, M.W.; Fortin, N.; Hall, J.; Ouellette, D.H.; Fraser, C.
 submitted to the EMBL Data Library, August 1994
 A;Description: The sequence of chromosome 1 of Saccharomyces cerevisiae.
 A;Reference number: S51956
 A;Accession: S51997

A;Molecule type: DNA
 A;Residues: 1-1435 <BUS>
 A;Cross-references: EMBL:U12980; NID:g1326053; PIDN:AA05008.1; PID:9595562; GSPDB:GN000148
 R;Keng, T.; Clark, M.W.; Storms, R.K.; Fortin, N.; Zhong, W.; Ouellette, B.F.F.; Barto
 Yeast 10, 953-958, 1994
 A;Title: LTE1 of Saccharomyces cerevisiae is a 1435 codon open reading frame that has
 A;Reference number: S45454; MUID:95076714; PMID:7985422
 A;Accession: S45454
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-301, 'C', 303-1435 <KEX>
 A;Cross-references: EMBL:L20125
 R;Shirayama, M.; Matsui, Y.; Tanaka, K.; Toh-E, A.
 Yeast 10, 451-461, 1994
 A;Title: Isolation of a CDC25 family gene, MS12/LTE1, as a multicopy suppressor of ira
 A;Reference number: S43456; MUID:95028143; PMID:7941731
 A;Accession: S43456
 A;Molecule type: DNA
 A;Residues: 1-997, 'LIVH', 1002, 'RKCIDN', 1010-1435 <SHI>
 A;Cross-references: GB:D21354; NID:G426455; PIDN:BA04620.1; PID:G452242
 R;Wickner, R.B.; Koh, T.J.; Crowley, J.C.; O'Neill, J.; Kaback, D.B.
 Yeast 3, 51-57, 1987
 A;Title: Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae: isolation
 A;Reference number: S05869; MUID:89073921; PMID:3332963
 A;Accession: S05869
 A;Molecule type: DNA
 A;Residues: 1127-1160, 'GB', 1164-1435 <WIC>
 A;Cross-references: EMBL:M16076; NID:G171849; PIDN:AAA34746.1; PID:G171850
 R;Keng, T.; Clark, M.W.; Storms, R.K.; Fortin, N.; Zhong, W.; Ouellette, B.F.F.; Barto
 submitted to the EMBL Data Library, December 1993
 A;Description: LTE1 of Saccharomyces cerevisiae is a 1435 codon open reading frame tha
 A;Reference number: S46646
 A;Accession: S46646
 A;Molecule type: DNA
 A;Residues: 1-1435 <KEX>
 A;Cross-references: EMBL:L20125; NID:G437022; PIDN:AAAS0468.1; PID:G437023
 C;Genetics:
 A;Gene: SGD:LTE1; MS12; MIPS:YAL024c
 A;Cross-references: SGD:S0000022; MIPS:YAL024c
 A;Map position: 1L
 C;Superfamily: guanine nucleotide-releasing factor LTE1; CDC25-type guanine nucleotide
 F;I190-1430/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match 96.3%; Score 26; DB 1; Length 1435;
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | ||||
 Db 690 PTLKTK 695

RESULT 10
 B87379
 ribosomal protein L35 [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 02-Aug-2002
 C;Accession: B87379
 R;Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: B87379
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-66 <STO>
 A;Cross-references: GB:AE005673; NID:G13422342; PIDN:AAK23030.1; GSPDB:GN00148
 C;Genetics:
 C;Superfamily: Escherichia coli ribosomal protein L35

Query Match 92.6%; Score 25; DB 2; Length 66;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 |
 |
 |
 |
 Db 2 PKLKTK 7

RESULT 11
 B71666
 ribosomal protein L35 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: B71666
 R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichertz-Ponten, T.; Alemark, U.
 Nature 396, 133-140, 1998
 A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: B71666
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Status: preliminary; DNA
 A:Molecule type: DNA
 A:Residues: 1-67 <AND>
 A:Cross-references: GB:AJ235272; GB:AJ235269; MID:93861033; PIDN:CAA15052.1; PID:9386115
 A:Experimental source: strain Madrid E
 C:Genetics:
 A:Gene: rpmI; RP608
 C:Superfamily: Escherichia coli ribosomal protein L35

Query Match 92.6%; Score 25; DB 2; Length 67;
 Best Local Similarity 83.3%; Pred. No. 24;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 |
 |
 |
 |
 Db 2 PKLKTK 7

RESULT 12
 F64423
 hypothetical protein homolog MJ0990 - Methanococcus jannaschii
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: F64423
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: F64423
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Status: preliminary; DNA
 A:Molecule type: DNA
 A:Residues: 1-75 <BUL>
 A:Cross-references: GB:U67541; GB:L77117; MID:92826353; PIDN:AAB98992.1; PID:91591652; T
 C:Genetics:
 A:Map position: REV919929-919702
 C:Superfamily: conserved hypothetical protein HI0721

Query Match 92.6%; Score 25; DB 2; Length 75;
 Best Local Similarity 83.3%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 |
 |
 |
 |
 Db 18 PVLKTK 23

RESULT 13
 F97816
 50S ribosomal protein L35 [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 17-May-2002
 C:Accession: F97816
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: F97816
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-89 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL03472.1; PID:915620044; GSPDB:GN00173
 C:Genetics:
 A:Gene: rpmI
 C:Superfamily: Escherichia coli ribosomal protein L35

Query Match 92.6%; Score 25; DB 2; Length 89;
 Best Local Similarity 83.3%; Pred. No. 33;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 |
 |
 |
 |
 Db 23 PKLKTK 28

RESULT 14
 C75271
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: C75271
 R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: C75271
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-142 <WHI>
 A:Cross-references: GB:AE002075; GB:AE000513; MID:96460272; PIDN:AAF12007.1; PID:96460
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2457
 A:Map position: 1

Query Match 92.6%; Score 25; DB 2; Length 142;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 |
 |
 |
 |
 Db 88 PVLKTK 93

RESULT 15
 T44368
 hypothetical protein [imported] - Staphylococcus aureus (fragment)
 C:Species: Staphylococcus aureus
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 11-May-2000
 C:Accession: T44368
 R:Kuroda, M.; Hayashi, H.; Ohta, T.
 Microbiol. Immunol. 43, 115-125, 1999
 A:Title: Chromosome-determined zinc responsible operon czr in Staphylococcus aureus
 A:Reference number: Z22754; MUID:99244271; PMID:10229265
 A:Accession: T44368
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-160 <KUR>
 A:Cross-references: EMBL:AB016431; PIDN:BAA36689.1
 A:Experimental source: strain 912

Query Match 92.6%; Score 25; DB 2; Length 160;

Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 PXLKTK 6
|
|
|
|
Db 7 PILKTK 12

Search completed: August 4, 2003, 16:08:00
Job time : 41 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:56:18 ; Search time 25 Seconds
(without alignments)
11.286 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLKTK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Lists various protein entries like Q00329 salmonella, P16371 drosophila, etc.

Table with columns: ID, RFBJ_SALMU, STANDARD, PRT, 293 AA. Lists protein IDs like 34, 35, 36, 37, etc.

ALIGNMENTS

RESULT 1
RFBJ_SALMU
AC Q00329: STANDARD; PRT; 293 AA.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE CDP-abequose synthase (EC 4.2.1.-)
GN RFBJ.
OS Salmonella muenchen.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=596;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M67.
RX MEDLINE=92349966; PubMed=1379320;
RA Brown P.K., Romana L.K., Reeves P.R.;
RT "Molecular analysis of the rfb gene cluster of Salmonella serovar
muenchen (strain M67): the genetic basis of the polymorphism between
groups C2 and B.";
RL Mol. Microbiol. 6:1385-1394(1992).
CC -!- CATALYTIC ACTIVITY: CDP-4-keto 3,6-dideoxyglucose = CDP-abequose.
CC -!- COFACTOR: NAD.
CC -!- PATHWAY: CDP-ABEQUOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.iab-sib.ch/announce/
or send an email to license@iab-sib.ch).
CC -----
DR EMBL; X61917; CAA43918.1;
KW Lipopolysaccharide biosynthesis; Lyase; NAD.
SQ SEQUENCE 293 AA; 33775 MW; F7E88187B2E87B17 CRC64;
Query Match 96.3%; Score 26; DB 1; Length 293;
Best Local Similarity 83.3%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PXLKTK 6
DB 103 PALKTK 108
RESULT 2
SPS4_YEAST STANDARD; PRT; 338 AA.
AC P09937;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Sporulation-specific protein 4.
DE

GN SPS4 OR YOR313C OR O6120.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=87089807; PubMed=3540611;
RX Carber A.T., Segall J.;
RT "The SPS4 gene of Saccharomyces cerevisiae encodes a major
RT sporulation-specific mRNA";
RT Mol. Cell. Biol. 6:4478-4485 (1986).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C / FY1679;
RX MEDLINE=97051589; PubMed=8896266;
RA Pearson B.M., Hernandez Y., Payne J., Wolf S.S., Kalogeropoulos A.,
RA Schweizer M.,
RT "Sequencing of a 35.71 kb DNA segment on the right arm of yeast
RT chromosome XV reveals regions of similarity to chromosomes I and
RT XIII";
RL Yeast 12:1021-1031 (1996).
CC -1- FUNCTION: NOT ESSENTIAL FOR SPORULATION. MIGHT BE A COMPONENT OF
CC THE CELL WALL.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT 6 OF 8 HOURS OF SPORULATION
CC WITH MAXIMAL TRANSCRIPT ACCUMULATION OCCURRING AT 8 TO 12 H, A
CC TIME AT WHICH THE MEIOTIC EVENTS IF SPORULATION HAVE BEEN
CC COMPLETED & THE DEPOSITION OF SPORE WALL COMPONENTS IS BEGINNING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; M14684; AAA35081.1; -
DR EMBL; X90565; CAA62168.1; -
DR EMBL; 275221; CAA99633.1; -
DR FIR; S58324; S58324.
DR SGD; S0005840; SPS4.
DR GO; GO:0006259; P:DNA metabolism; IGI.
DR GO; GO:0007126; P:meiosis; IEP.
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); IEP.
KW Sporulation; Meiosis
FT CONFLICT 107 107 R -> H (IN REF. 1).
FT CONFLICT 180 180 H -> D (IN REF. 1).
FT SEQUENCE 338 AA; 38591 MW; 253C38A9D43F07F2 CRC64;
QY 1 PMLKTK 6
Db 135 PSLKTK 140
Query Match 96.3%; Score 26; DB 1; Length 338;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ID GROU_DROME STANDARD; PRT; 719 AA.
AC F16371; Q9V3F7;
DT 01-AUG-1990 (Rel. 15, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Groucho protein (Enhancer of split M9/10).
GN GRK OR F(SPL)M9/M10 OR BCDNA:ID33829 OR CG8384.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN SEQUENCE FROM N.A.
RX MEDLINE=89051868; PubMed=3142687;
RA Hartley D., Preiss A., Artavanis-Tsakonas S.;
RT "A deduced gene product from the Drosophila neurogenic locus,
RT enhancer of split, shows homology to mammalian G-protein beta
RT subunit";
RT Cell 55:785-795 (1988).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeung M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Balgwin D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch L., Beasley E.M.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Bozhatov S.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Garg N.S., Gelbart W.M., Glasser K.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.-H., Wang X.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang J.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [3]
RN SEQUENCE FROM N.A.
RP STRAIN=Oregon-R; TISSUE=Embryo;
RX MEDLINE=20196012; PubMed=10731138;
RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
RA Stapleton M., Harvey D.A.;
RT "A Drosophila complementary DNA resource";
RL Science 287:2222-2224 (2000).
RN [4]
RN BINDING TO HAIRY-RELATED PROTEINS.
RX MEDLINE=95094252; PubMed=8001118;
RA Paroush Z., Finley R.L. Jr., Kidd T., Wainwright S.M., Ingham P.W.,
RA Brent R., Ieh-Horowitz D.;
RT "Groucho is required for Drosophila neurogenesis, segmentation, and
RT sex determination and interacts directly with hairy-related bHLH
RT proteins";
RL Cell 79:805-815 (1994).
CC -1- FUNCTION: INVOLVED IN NEUROGENESIS; IN THE SEGREGATION OF THE
CC NEUROCYTODERM. DIRECTLY OR INDIRECTLY INTERACTS WITH NOTCH AND
CC DELTA.
CC -1- SUBUNIT: FORMS A COMPLEX WITH THE HAIRY/ENHANCER OF SPLIT/DEADPAN
CC FAMILY OF BASIC HELIX-LOOP-HELIX PROTEINS IN ORDER TO REPRESS
CC TRANSCRIPTION. ITS ACTIVITY IN REGULATING TRANSCRIPTION DEPENDS ON

CC OTHER PROTEINS AS IT LACKS A DNA-BINDING MOTIF.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 6 WD repeats.
 CC -1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M20571; AAA28512.1; -
 DR EMBL; AE003754; NAF56556.1; -
 DR EMBL; AF145695; AAD38670.1; -
 DR TRANSFAC; T02451; -
 DR FLYBase; FBgn0001139; gro.
 DR GO; GO:0003714; P:transcription co-repressor activity; IPI.
 DR GO; GO:0045810; P:negative regulation of frizzled receptor sig. . . ; NAS.
 DR GO; GO:0007399; P:neurogenesis; IMP.
 DR InterPro; IPR005617; TLE_N.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF031920; TLE_N; 1.
 DR Pfam; PF00400; WD40; 6.
 DR ProDom; PD000018; WD40; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD REPEATS_1; 2.
 DR PROSITE; PS50082; WD REPEATS_2; 3.
 DR PROSITE; PS50294; WD REPEATS_REGION; 2.
 KW Differentiation; Neurogenesis; Nuclear protein; Repeat; WD repeat;
 KW Phosphorylation; Wnt signaling pathway; Transcription regulation.
 FT DOMAIN 1 133
 FT DOMAIN 134 194
 FT DOMAIN 195 256
 FT DOMAIN 257 398
 FT BINDING 251 414
 FT DOMAIN 257 398
 FT REPEAT 431 460
 FT REPEAT 477 507
 FT REPEAT 521 551
 FT REPEAT 563 593
 FT REPEAT 645 675
 FT REPEAT 716
 FT MOD RES 231 231
 FT MOD RES 247 247
 FT CONFLICT 41 41
 SQ SEQUENCE 719 AA; 78919 MW; D495291FD77026A5 CRC64;

Query Match 96.3%; Score 26; DB 1; Length 719;
 Best Local Similarity 83.3%; Pred. No. 60;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 PXLKTK 6
 DB 296 PSLKTK 301
 |||||
 |||||

RESULT 4
 ID _TLE1_YEAST STANDARD; PRT; 1435 AA.
 AC P07866;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Low temperature essential protein.
 GN _TLE1 OR MSI2 OR YAL024C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;

RN SEQUENCE FROM N.A.
 RP STRAIN=S288C / AB972;
 RX MEDLINE=95076714; PubMed=7985422;
 RA Keng T., Clark M.W., Storms R.K., Fortin N., Zhong W.,
 RA Ouellette F.B.F., Barton A.B., Kaback D.B., Bussey H.;
 RT "L1 of Saccharomyces cerevisiae is a 1435 codon open reading frame
 RT that has sequence similarities to guanine nucleotide releasing
 RT factors."
 RL Yeast 10:953-958(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
 RA Hall J., Ouellette F.B.F., Keng T., Barton A.B., Su Y., Davies C.K.,
 RA Storms R.K.;
 RT "The nucleotide sequence of chromosome I from Saccharomyces
 RT cerevisiae."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95028143; PubMed=7941731;
 RA Shirayama M., Matsui Y., Tanaka K., Toh-E A.;
 RT "Isolation of a CDC25 family gene, MSI2/LTE1, as a multicopy
 RT suppressor of *ira1*."
 RL Yeast 10:451-461(1994).
 RN [4]
 RP SEQUENCE OF 1127-1435 FROM N.A.
 RX MEDLINE=89073921; PubMed=3332963;
 RA Wickner R.B., Koh T.J., Crowley J.C., O'Neil J., Kaback D.B.;
 RT "Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae:
 RT isolation of the MAK16 gene and analysis of an adjacent gene
 RT essential for growth at low temperatures."
 RL Yeast 3:51-57(1987).
 RN [5]
 RP SEQUENCE OF 1404-1435 FROM N.A.
 RX MEDLINE=88320371; PubMed=3045810;
 RA Wickner R.B.;
 RT "Host function of MAK16: G1 arrest by a mak16 mutant of Saccharomyces
 RT cerevisiae."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:6007-6011(1988).
 CC -1- FUNCTION: PUTATIVE GDP-GTP EXCHANGE FACTOR FOR A RAS-LIKE PROTEIN.
 CC THIS PROTEIN IS ESSENTIAL FOR GROWTH AT LOW TEMPERATURES.
 CC INVOLVED IN THE TERMINATION OF M PHASE.
 CC -1- SIMILARITY: Contains 1 Ras-GEF domain.
 CC -1- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L20125; AAA50468.1; -
 DR EMBL; U12980; AAC05008.1; -
 DR EMBL; D21354; BAA04820.1; -
 DR EMBL; M16076; AAA34746.1; -
 DR EMBL; J03852; AAA34751.1; -
 DR PIR; S51997; BVBYL1.
 DR SGD; S0000022; LTE1.
 DR InterPro; IPR000651; RasGEFN.
 DR InterPro; IPR001895; RasGRF_CDC25.
 DR Pfam; PF00617; RasGEF; 1.
 DR Pfam; PF00618; RasGEFN; 1.
 DR SMART; SM00147; RasGEF; 1.
 DR SMART; SM00229; RasGEFN; 1.
 DR PROSITE; PS00720; RasGEF; 1.
 DR PROSITE; PS50009; RasGEF_CAT; 1.
 DR PROSITE; PS50212; RasGEF_NTER; 1.
 DR Guanine-nucleotide releasing factor.

FT DOMAIN 25 157 N-TERMINAL RAS-GEF.
 FT DOMAIN 1194 1434 RAS-GEF.
 FT CONFLICT 998 1009 TNSNIGSVLTM -> LIVHIRKIDN (IN REF. 3).
 FT CONFLICT 1161 1163 AAQ -> GE (IN REF. 4).
 SQ SEQUENCE 1435 AA, 163149 MW, EED7E5150BECA3DE CRC64;
 Query Match 96.3%; Score 26; DB 1; Length 1435;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 690 PTLKTK 695

RESULT 5
 RL35_CAUCR STANDARD; PRT; 66 AA.
 ID_RL35_CAUCR
 AC Q9A9E2; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L35.
 GN RPM1 OR CC1046.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Rickettsiales; Rickettsiaceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RC SEQUENCE FROM N.A.
 PC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nerlan W.C.; Feldblyum T.V.; Laub M.T.; Paulsen I.T.; Nelson K.E.,
 Eisen J., Heideberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 Potocha I., Nelson W.C., Durkin A., Stephens C., Phadke N.D., Ely B.,
 Ra DeBoy R.T., Dodson R.J., Newton A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 Uterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4138-4141(2001).
 CC -!- SIMILARITY: BELONGS TO THE L35P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ235272; CAAL5052.1; -
 DR PIR; B71666; B71666.
 DR HAMAP; MF 00514; -; 1.
 DR InterPro; IPR001706; Ribosomal_L35.
 DR Pfam; PF01632; Ribosomal_L35p; 1.
 DR PRINTS; PR00064; RIBOSOMAL_L35; 1.
 DR PRODom; PD003417; Ribosomal_L35; 1.
 DR TIGRFAMs; TIGR00001; rpm1_bact; 1.
 DR TRIPLEX; TRIPLEX001; RIBOSOMAL_L35; 1.
 DR PROSITE; PS00936; RIBOSOMAL_L35; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 66 AA; 7354 MW; 0C0715ED3B575BBB CRC64;

Query Match 92.6%; Score 25; DB 1; Length 66;
 Best Local Similarity 83.3%; Pred. No. 9.9;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 690 PTLKTK 695

RESULT 6
 RL35_RICPR STANDARD; PRT; 67 AA.
 ID_RL35_RICPR
 AC Q9ZCVL; 2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L35.
 GN RPM1 OR RP608.
 GN RPM1 OR RP608.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiales; Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RC SEQUENCE FROM N.A.
 PC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sichteritz-Ponten T., Alismark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria."
 RL Nature 396:133-140(1998).
 CC -!- SIMILARITY: BELONGS TO THE L35P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ235272; CAAL5052.1; -
 DR PIR; B71666; B71666.
 DR HAMAP; MF 00514; -; 1.
 DR InterPro; IPR001706; Ribosomal_L35.
 DR Pfam; PF01632; Ribosomal_L35p; 1.
 DR PRINTS; PR00064; RIBOSOMAL_L35; 1.
 DR PRODom; PD003417; Ribosomal_L35; 1.
 DR TIGRFAMs; TIGR00001; rpm1_bact; 1.
 DR TRIPLEX; TRIPLEX001; RIBOSOMAL_L35; 1.
 DR PROSITE; PS00936; RIBOSOMAL_L35; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 67 AA; 7705 MW; 3CF2F9F4E4BF5BA1 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 67;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 690 PTLKTK 695

RESULT 7
 RL35_RICCN STANDARD; PRT; 68 AA.
 ID_RL35_RICCN
 AC Q92H8; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 50S ribosomal protein L35.
 GN RPM1 OR RC0934.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiales; Rickettsiaceae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RC SEQUENCE FROM N.A.
 PC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renest-Oudiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,

Query Match 92.6%; Score 25; DB 1; Length 67;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 Db 690 PTLKTK 695

RA Raoult D.;
 RL "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001);
 CC -!- SIMILARITY: BELONGS TO THE L35P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR HAMB; AE008647; AAL03472.1; ALT_INIT.
 DR HAMAP; MF_00514; 1
 DR InterPro; IPR001706; Ribosomal_L35.
 DR Pfam; PF01632; Ribosomal_L35p; 1.
 DR PRINTS; PR00064; RIBOSOMAL_L35.
 DR PRODOM; PD003417; Ribosomal_L35; 1.
 DR TIGRFAMs; TIGR00001; rpm1_bact; 1.
 DR PROSITE; PS00936; RIBOSOMAL_L35; 1.
 KW Ribosomal protein; Complete proteome.
 SQ SEQUENCE 68 AA; 7896 MW; DIA9A2F9E90AD3E0 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 68;
 Best Local Similarity 83.3%; Pred. No. 10;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | | | | |
 Db 2 PXLKTK 7

RESULT 8

Y990 METJA ID Y990 METJA STANDARD; PRT; 75 AA.
 AC Q58357;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ0990.
 GN MJ0990.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcales; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAINS-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kervilave A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -!- SIMILARITY: BELONGS TO THE UPF0033 FAMILY.
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U67541; AAB98992.1;
 DR PIR; F64423; F64423.

DR TIGR; MJ0990; -;
 DR InterPro; IPR001455; UPF00033.
 DR Pfam; PF01206; UPF0033; 1.
 DR PROSITE; PS01148; UPF0033; 1.
 KW Hypothetical protein; Complete proteome.
 FT DOMAIN 27 36 GLU-RICH.
 SQ SEQUENCE 75 AA; 8465 MW; CF7009364C388539 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 75;
 Best Local Similarity 83.3%; Pred. No. 11;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | | | | |
 Db 18 PVLKTK 23

RESULT 9

RL5 MICLU ID RL5 MICLU STANDARD; PRT; 191 AA.
 AC P33098;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-1993 (Rel. 27, Last annotation update)
 DE 50S ribosomal protein L5.
 GN RPL5.
 OS Micrococcus luteus (Micrococcus lysodeikticus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Micrococcaceae; Micrococcus.
 OX NCBI_TaxID=1270;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90133967; PubMed=2533272;
 RA Ohama T., Mato A., Osawa S.;
 RT "Spectinomycin operon of Micrococcus luteus: evolutionary
 RT implications of organization and novel codon usage.";
 RL J. Mol. Evol. 29:381-395(1989).
 CC -!- FUNCTION: THIS IS ONE OF 3 PROTEINS THAT MEDIATE THE ATTACHMENT OF
 CC THE 5S RNA INTO THE LARGE RIBOSOMAL SUBUNIT (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE L5P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X17524; CAA35560.1; -;
 DR PIR; S29884; S29884.
 DR InterPro; IPR002132; Ribosomal_L5.
 DR InterPro; IPR003236; Ribosomal_L5_mit.
 DR Pfam; PF00281; Ribosomal_L5; 1.
 DR Pfam; PF00673; Ribosomal_L5_C; 1.
 DR PRODOM; PD001076; Ribosomal_L5; 1.
 DR PRODOM; PD013434; Ribosomal_L5_mit; 1.
 DR PROSITE; PS00358; RIBOSOMAL_L5; 1.
 KW Ribosomal protein; rRNA-binding.
 SQ SEQUENCE 191 AA; 21674 MW; 2B98E58FF5F2505F CRC64;

Query Match 92.6%; Score 25; DB 1; Length 191;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 | | | | |
 Db 12 PRLKTK 17

RESULT 10

SET1_CABEL ID SET1_CABEL STANDARD; PRT; 242 AA.

15-SEP-2003 (Rel. 42, Last annotation update)
 Oligoribonuclease, mitochondrial precursor (EC 3.1.1.-.-).
 RX2 OR YNT20 OR YLR059C OR L2159.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=S288C;
 RA Hanekamp T., Thorsness P.E.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 SEQUENCE FROM N.A.
 STRAIN=S288C / AB972;
 RC MEDLINE=97313267; PubMed=9169871;
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,
 Benes V., Brueckner M., Delius H., Dubois E., Duesterhoeft A.,
 Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
 Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
 Louis E.J., Messing F., Mewes H.-W., Miosga T., Moesli D.,
 Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
 Portetle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
 Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
 Underwood A.P., Urrestazu L.A., Vandebol M., Verhasselt P.,
 Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
 Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
 "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
 Nature 387:87-90(1997).
 [3]
 IDENTIFICATION AND SUBCELLULAR LOCATION.
 RP MEDLINE=99132184; PubMed=9933355;
 RA Hanekamp T., Thorsness P.E.;
 RT "YNT20, a bypass suppressor of yme2, encodes a putative 3'-5'
 exonuclease localized in mitochondria of Saccharomyces cerevisiae.";
 Curr. Genet. 34:438-448(1999).
 RL FUNCTION: 3'-TO-5' EXORIBONUCLEASE SPECIFIC FOR SMALL
 CC OLIGORIBONUCLEOTIDES (BY SIMILARITY).
 CC SUBCELLULAR LOCATION: Mitochondrial.
 CC SIMILARITY: BELONGS TO THE OLIGORIBONUCLEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; L47977; AAA98633.1; -;
 DR EMBL; X94607; CAA64306.1; -;
 DR EMBL; Z73231; CAA97590.1; -;
 DR PIR; S61633; S61633.
 DR SGD; S0004049; REX2.
 DR GO; GO:0005739; C:mitochondrion; IDA.
 DR GO; GO:0008408; F:3'-5' exonuclease activity; IMP.
 DR GO; GO:0006396; P:RNA processing; IMP.
 DR InterPro; IPR006055; Exonuclease.
 DR Pfam; PF00929; Exonuclease; 1.
 DR SMART; SM00479; EXOIII; 1.
 DR KW HydroLase; Exonuclease; Nuclease; Mitochondrion; Transit peptide.
 FT TRANSIT ?
 FT CHAIN ?
 FT ACT SITE 184 184
 FT CONFLICT 100 100
 FT CONFLICT 124 127
 FT CONFLICT 223 269
 FT REF. 1)
 FT S -> T (IN REF. 1).
 FT GLTA -> ESHP (IN REF. 1).
 FT SIAOLQWYMDVNLKPKQETESVIGSEQEPSPSSTSLK
 FT RPTDF -> HSSIAMVHQLELLEATTGNRVGRVRI (IN
 FT REF. 1).
 FT SEQUENCE 269 AA; 30870 MW; C942C823691E9815 CRC64;
 Query Match 92.6%; Score 25; DB 1; Length 269;
 Best Local Similarity 83.3%; Pred. No. 41;

DT 022795;
 DE 15-SEP-2003 (Rel. 42, Created)
 GN 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein set-1.
 GN SET-1 OR T26A5.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
 Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=Bristol N2;
 RC Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RA Du Z.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 REVISIONS.
 RA Waterston R.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 SUBCELLULAR LOCALIZATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RP MEDLINE=22114964; PubMed=12119097;
 RA Texanova R., Pujol N., Fasano L., Djabali M.;
 RT "Characterisation of set-1, a conserved PR/SET domain gene in
 Caenorhabditis elegans.";
 Gene 292:33-41(2002).
 RL FUNCTION: Essential protein probably involved in chromatin
 modification and/or regulation.
 CC SUBCELLULAR LOCATION: Nuclear.
 CC TISSUE SPECIFICITY: In embryos, it is expressed ubiquitously. In
 late embryos, it is expressed in hypodermal seam cells. In L3 and
 L4 larvae and thereafter, it is expressed in vulval precursor
 cells. In adult males, it is also expressed in 6 unidentified
 posterior cells.
 CC DEVELOPMENTAL STAGE: Highly expressed in eggs, then decreases.
 CC SIMILARITY: Contains 1 SET domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 DR EMBL; U00043; AAC77512.1; -;
 DR PIR; T34384; T34384.
 DR WormPep; T26A5.7; CE19602.
 DR InterPro; IPR001214; SET.
 DR Pfam; PF00856; SET; 1.
 DR SMART; SM00317; SET; 1.
 DR PROSITE; PPS0280; SET; 1.
 DR KW Nuclear protein; Developmental protein.
 FT DOMAIN 103 230
 FT POLY-ALA.
 FT DOMAIN 18 21
 FT POLY-SER.
 FT DOMAIN 38 43
 FT SEQUENCE 242 AA; 27568 MW; 0F752B79505AFA99 CRC64;
 Query Match 92.6%; Score 25; DB 1; Length 242;
 Best Local Similarity 83.3%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PXLKTK 6
 Db 192 PNLKTK 197
 RESULT 11
 ORN YEAST STANDARD; PRT; 269 AA.
 AC P54964;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 | | | | |
 Db 45 PELKTK 50

RESULT 12
 AEPE_BPA18 STANDARD; PRT; 281 AA.
 AC Q37976; Q9T199;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE L-alanyl-D-glutamate peptidase (EC 3.4.-.-).
 GN PLY OR PLY18.
 OS Bacteriophage A118.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=40521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96020653; PubMed=8577256;
 RA Loessner M.J., Wendlinger G., Scherer S.;
 RT "Heterogeneous endolysins in *Listeria monocytogenes* bacteriophages: a
 RT new class of enzymes and evidence for conserved holin genes within
 RT the siphoviral lysis cassettes.";
 RL Mol. Microbiol. 16:1231-1241(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20117992; PubMed=10652093;
 RA Loessner M.J., Imman R.B., Lauer P., Calendar R.;
 RT "Complete nucleotide sequence, molecular analysis and genome structure
 RT of bacteriophage A118 of *Listeria monocytogenes*: implications for
 RT phage evolution.";
 RL Mol. Microbiol. 35:324-340(2000).
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes the link between L-alanine and
 CC D-glutamate residues in certain bacterial cell-wall glycopeptides.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ABOUT 20 MINUTES AFTER
 CC INFECTION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X85008; CAA59362.1; --
 DR EMBL; AJ242593; CAB53811.1; --
 DR PIR; S69799; S69799.
 DR MEROPS; M15_020; --
 DR InterPro; IPR003709; VanY.
 DR Pfam; PF02557; VanY; 1.
 KW Hydrolyase; Cell wall.
 FT CONFLICT 24 24 A -> R (IN REF. 1).
 SQ SEQUENCE 281 AA; 30799 MW; 9CADC9F02F54CB41 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 281;
 Best Local Similarity 83.3%; Pred. No. 42;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 | | | | |
 Db 261 PELKTK 266

RESULT 13
 ALC_YEAST STANDARD; PRT; 343 AA.
 AC P25335;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Allantoicase (EC 3.5.3.4) (Allantoate amidohydrolyase).
 GN DAL2 OR ALCI OR YI029W
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92009196; PubMed=19162777;
 RA Yoo H.S., Cooper T.G.;
 RT "Sequences of two adjacent genes, one (DAL2) encoding allantoicase
 RT and another (DCG1) sensitive to nitrogen-catabolite repression in
 RT Saccharomyces cerevisiae.";
 RL Gene 104:55-62(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92206079; PubMed=1839481;
 RA Lee F.-J.S., Moss J.;
 RT "Cloning of a Saccharomyces cerevisiae gene encoding a protein
 RT homologous to allantoicase of *Neurospora crassa*.";
 RL Yeast 7:993-995(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
 RA Harris D.E., Hornsell T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
 RL Nature 387:84-87(1997).
 CC -1- FUNCTION: UTILIZATION OF PURINES AS SECONDARY NITROGEN SOURCES,
 CC WHEN PRIMARY SOURCES ARE LIMITING.
 CC -1- CATALYTIC ACTIVITY: Allantoate + H(2)O = (-)-ureidoglycolate +
 CC urea.
 CC -1- PATHWAY: Degradation of allantoin (purine catabolism); second
 CC step.
 CC -1- INDUCTION: REPRESSED BY NITROGEN.
 CC -1- SIMILARITY: BELONGS TO THE ALLANTOICASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M64720; AAA34554.1; --
 DR EMBL; X60460; CAA42994.1; --
 DR EMBL; Z38061; CAA86189.1; --
 DR PIR; JH0442; JH0442.
 DR SGD; S0001468; DAL2.
 DR GO; GO:0000256; P:allantoin catabolism; IMP.
 DR InterPro; IPR005164; Allantoicase.
 DR Pfam; PF03561; Allantoicase; 2.
 DR Hydrolyase; Purine metabolism.
 KW CONFLICT 93 93 A -> S (IN REF. 2).
 FT CONFLICT 134 135 WV -> SL (IN REF. 2).
 SQ SEQUENCE 343 AA; 38714 MW; 0F9CB0FBA5EB76F1 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 343;
 Best Local Similarity 83.3%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 | | | | |
 Db 187 PELKTK 192

RESULT 14

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBITaxID=9606;
RN [1]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBITaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Renal cell carcinoma;
RX MEDLINE=99438124; PubMed=10508479;
RA Scanlan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H.,
RA Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA Old L.J.;
RT "Antigens recognized by autologous antibody in patients with renal-
RT cell carcinoma.";
RL Int. J. Cancer 83:456-464 (1999).
RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE=Renal cell carcinoma;
RX MEDLINE=99438124; PubMed=10508479;
RA Scanlan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H.,
RA Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA Old L.J.;
RT "Antigens recognized by autologous antibody in patients with renal-
RT cell carcinoma.";
RL Int. J. Cancer 83:456-464 (1999).
RN [2]

RP SEQUENCE FROM N.A.
RC Roberts T.P., Wright A., Wahab N.A., Weston B.S., Mason R.M.;
RT "Gene which is selectively expressed in hyperglycaemia.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 YTH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF155095; AAD42861.1; --
DR EMBL; AF155095; AAD42861.1; --
DR GO; GO:0008222; P:tumor antigen; TAS.
DR GO; GO:0006959; P:humoral immune response; TAS.
DR Pfam; PF04146; YTH; 1.
DR PROSITE; PS50882; YTH; 1.
FT DOMAIN 410 544 YTH.
SQ SEQUENCE 570 AA; 61320 MW; 91FCFAYE508869E4 CRC64;

RP SEQUENCE FROM N.A.
RC Roberts T.P., Wright A., Wahab N.A., Weston B.S., Mason R.M.;
RT "Gene which is selectively expressed in hyperglycaemia.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 YTH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF155095; AAD42861.1; --
DR EMBL; AF155095; AAD42861.1; --
DR GO; GO:0008222; P:tumor antigen; TAS.
DR GO; GO:0006959; P:humoral immune response; TAS.
DR Pfam; PF04146; YTH; 1.
DR PROSITE; PS50882; YTH; 1.
FT DOMAIN 410 544 YTH.
SQ SEQUENCE 570 AA; 61320 MW; 91FCFAYE508869E4 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 570;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PXLKTK 6
| | | | |
Db 251 PDLKTK 283

Query Match 92.6%; Score 25; DB 1; Length 570;
Best Local Similarity 83.3%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 PXLKTK 6
| | | | |
Db 251 PDLKTK 283

Search completed: August 4, 2003, 16:05:30
Job time : 26 secs

Search completed: August 4, 2003, 16:05:30
Job time : 26 secs

PGK_STRCO STANDARD; PRT; 403 AA.
AC Q9Z519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR SCO1946 OR SCC54.06C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBITaxID=1902;
RN [1]

PGK_STRCO STANDARD; PRT; 403 AA.
AC Q9Z519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR SCO1946 OR SCC54.06C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBITaxID=1902;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -!- PATHWAY: Second phase of glycolysis; second step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL035591; CAB38136.1; --
DR EMBL; AL035591; CAB38136.1; --
DR PIR; T36019; T36019.
DR HSSP; P18912; 1PHP.
DR HAMAP; MF_00145; ; 1.
DR InterPro; IPR001576; ; PGK.
DR Pfam; PF00162; PGK; 1.
DR PRINTS; PR00477; PHGLYCKINASE.
DR PROSITE; PS00111; PGLYCERATE KINASE; 1.
DR Transferase; Kinase; Glycolysis; Complete proteome.
KW SEQUENCE 403 AA; 41765 MW; C42094E7C6221FE5 CRC64;

RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -!- PATHWAY: Second phase of glycolysis; second step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phosphoglycerate kinase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL035591; CAB38136.1; --
DR EMBL; AL035591; CAB38136.1; --
DR PIR; T36019; T36019.
DR HSSP; P18912; 1PHP.
DR HAMAP; MF_00145; ; 1.
DR InterPro; IPR001576; ; PGK.
DR Pfam; PF00162; PGK; 1.
DR PRINTS; PR00477; PHGLYCKINASE.
DR PROSITE; PS00111; PGLYCERATE KINASE; 1.
DR Transferase; Kinase; Glycolysis; Complete proteome.
KW SEQUENCE 403 AA; 41765 MW; C42094E7C6221FE5 CRC64;

Query Match 92.6%; Score 25; DB 1; Length 403;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
QY 1 PXLKTK 6
| | | | |
Db 278 PDLKTK 283

Query Match 92.6%; Score 25; DB 1; Length 403;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
QY 1 PXLKTK 6
| | | | |
Db 278 PDLKTK 283

RESULT 15
HGR8_HUMAN STANDARD; PRT; 570 AA.
AC Q9Y519;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE High-glucose-regulated protein 8 (NY-REN-2 antigen).
GN HGRG8.
OS Homo sapiens (Human).

RESULT 15
HGR8_HUMAN STANDARD; PRT; 570 AA.
AC Q9Y519;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE High-glucose-regulated protein 8 (NY-REN-2 antigen).
GN HGRG8.
OS Homo sapiens (Human).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 15:56:48 ; Search time 96 Seconds
(without alignments)
16.128 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLTKK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp Vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	26	96.3	128	3 Q96VM5	Q96vm5 fusarium ox
2	26	96.3	138	10 O48890	O48890 arabiGopsis
3	26	96.3	242	12 Q84406	Q84406 paramacium
4	26	96.3	266	4 Q8TAX0	Q8tax0 homo sapien
5	26	96.3	275	11 Q8BNN4	Q8bnn4 mus muscucu
6	26	96.3	318	2 Q9R7C6	Q9r7c6 wolbachia p
7	26	96.3	335	16 Q8F7X4	Q8f7x4 leptospira
8	26	96.3	351	4 Q9H7B0	Q9h7b0 homo sapien
9	26	96.3	351	4 Q9BFP3	Q9bfp3 homo sapien
10	26	96.3	351	11 Q8C569	Q8c569 mus muscucu
11	26	96.3	446	3 Q96VM9	Q96vm9 fusarium ox
12	26	96.3	542	3 Q00832	Q00832 fusarium ox
13	26	96.3	568	12 Q9PYX6	Q9pyx6 xestia c-ni
14	26	96.3	577	4 Q96A20	Q96a20 homo sapien
15	26	96.3	577	6 Q9BEA2	Q9bea2 bos taurus
16	26	96.3	577	6 Q9TVB5	Q9tvb5 bos taurus

17	26	96.3	680	12 Q8V7I4	Q8v7i4 tt virus. o
18	26	96.3	732	12 Q9DUC7	Q9duc7 tt virus. o
19	26	96.3	738	12 Q8V7H2	Q8v7h2 tt virus. o
20	26	96.3	746	12 Q8V7H8	Q8v7h8 tt virus. o
21	26	96.3	749	5 O45381	O45381 caenorhabdi
22	26	96.3	774	5 Q9XZ53	Q9xz53 drosophila
23	26	96.3	899	10 P93698	P93698 vigna ungui
24	26	96.3	945	5 Q9UJ67	Q9uj67 caenorhabdi
25	26	96.3	979	13 O93509	O93509 xenopus lae
26	26	96.3	1695	5 Q9BKL2	Q9bkl2 hydra atten
27	26	96.3	2309	3 Q8WZV3	Q8wzv3 neurospora
28	26	96.3	2603	6 Q9XSS3	Q9xss3 potorous tr
29	26	96.3	3336	5 Q8IKE1	Q8ike1 plasmodium
30	25	92.6	51	4 Q9BXH2	Q9bxh2 homo sapien
31	25	92.6	51	11 Q99ML3	Q99ml3 mus muscucu
32	25	92.6	78	16 Q8KE43	Q8ke43 chlorobium
33	25	92.6	142	16 Q9RRN1	Q9rrn1 deinococcus
34	25	92.6	150	17 Q8PWF0	Q8pwf0 methanosarc
35	25	92.6	160	2 Q9ZNF3	Q9znf3 staphylococ
36	25	92.6	179	4 Q8WYH3	Q8wyh3 homo sapien
37	25	92.6	185	13 Q8JJ68	Q8jj68 oncorhynchu
38	25	92.6	185	16 Q9L0C8	Q9l0c8 streptomyce
39	25	92.6	193	4 Q9H8Q0	Q9h8q0 homo sapien
40	25	92.6	208	2 Q8KTG2	Q8ktg2 tomato big
41	25	92.6	213	2 O52279	O52279 agrobacteri
42	25	92.6	224	4 Q9BQ57	Q9bq57 homo sapien
43	25	92.6	231	16 O45543	O45543 bacillus su
44	25	92.6	242	5 Q22795	Q22795 caenorhabdi
45	25	92.6	250	16 Q985U5	Q985u5 rhizobium 1

ALIGNMENTS

RESULT 1
Q96VM5 PRELIMINARY; PRT; 128 AA.
ID Q96VM5;
AC Q96VM5; (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative transposase.
OS Fusarium oxysporum f. sp. ciceris.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OX Hypocreales; mitosporic Hypocreales; Fusarium.
OC NCBI_TaxID=62683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8905; TRANSPOSON=Potci;
RA Horman S.R., Bainbridge B.W.;
RT "Potci, a hAT family transposable element in Fusarium oxysporum f. sp.
RT ciceris";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY039814; AAK82933.1; -
DR InterPro; IPR006600; CENPB.
DR SMART; SM00674; CENPB; 1.
SQ SEQUENCE 128 AA; 14374 MW; 44742CF7F0B3D03A CRC64;

Query Match 96.3%; Score 26; DB 3; Length 128;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLTKK 6
DB 113 PALTKK 118

RESULT 2
O48890 PRELIMINARY; PRT; 138 AA.
ID O48890
AC O48890;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

RX MEDLINE=95407089; PubMed=7676624;
 RA Li Y., Lu Z., Burbank D.E., Kutish G.F., Rock D.L., Van Etten J.L.;
 RT "Analysis of 43 kb of the Chlorella virus PBCV-1 330-kb genome: map
 positions 45 to 88."
 RL Virology 212:134-150(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20013326; PubMed=10544099;
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., King W.,
 RA Lisec A.D., Nickerson K.W., Van Etten J.L.;
 RT "Chlorella virus PBCV-1 encodes a functional homosperrimidine
 synthase."
 RL Virology 263:254-262(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20478054; PubMed=11021991;
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
 PBCV-1."
 RL Virology 276:27-36(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Van Etten J.L.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RA Graves M.V., Van Etten J.L.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RA Gurnon J.R., Graves M.V., Van Etten J.L.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U42580; AAC96453.1; -;
 DR InterPro; IPR005123; 2OG-FeII_Oxy.
 DR InterPro; IPR006620; Pro_4_hyd_alph.
 DR Pfam; PF03171; 2OG-FeII_Oxy; 1.
 DR SMART; SM00702; P4HC; 1.
 SQ SEQUENCE 242 AA; 27812 MW; 3BFE77C5A1A6431A CRC64;

Query Match 96.3%; Score 26; DB 12; Length 242;
 Best Local Similarity 83.3%; Pred. No. 95;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 189 PTLKTK 194

RESULT 4
 Q8TAXO PRELIMINARY; PRT; 266 AA.
 ID Q8TAXO
 AC Q8TAXO;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to odd-skipped related 1 (Drosophila) (Zinc finger
 transcription factor) (Hypothetical protein FLJ90110).
 GN OSR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 15.9 kDa subunit of RNA polymerase II.
 GN RPB15.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Larkin R.M., Guilfoyle T.J.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=99156233; PubMed=10048489;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,091,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:379-391(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation."
 RL Genome Biol. 0:0-0(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Brewer V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF016511; AAB95261.1; -;
 DR EMBL; AB016893; BAB09413.1; -;
 DR EMBL; AY086086; AAM63291.1; -;
 DR InterPro; IPR005574; RNA_pol_Rpb4.
 DR Pfam; PF03874; RNA_pol_Rpb4; 1.
 SQ SEQUENCE 138 AA; 15932 MW; 337D7290D7C5BE31 CRC64;

Query Match 96.3%; Score 26; DB 10; Length 138;
 Best Local Similarity 83.3%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 110 PSLKTK 115

RESULT 3
 Q84406 PRELIMINARY; PRT; 242 AA.
 ID Q84406
 AC Q84406;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PBCV-1 prelyl 4-hydroxylase.
 GN A85R.
 OS Paramoecium bursaria chlorella virus 1 (PBCV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
 OX NCBI_TaxID=10506;
 RN [1]
 RP SEQUENCE OF 1-210 FROM N.A.
 RX MEDLINE=95133167; PubMed=7831789;
 RA Lu Z., Li Y., Zhang Y., Kutish G.F., Rock D.L., Van Etten J.L.;
 RT "Analysis of 45 kb of DNA located at the left end of the chlorella
 RT virus PBCV-1 genome."
 RL Virology 206:339-352(1995).
 RN [2]
 RP SEQUENCE OF 210-242 FROM N.A.

OX NCBI_TaxID=9606;
 RP [1] PXLKTK 6
 RC SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22115036; PubMed=12119563;
 RA Katch M.;
 RT "Molecular cloning and characterization of OSR1 on human chromosome
 RT 2p24."
 RL Int. J. Mol. Med. 10:221-225 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
 RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
 RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
 RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
 RT "NEDO human cdna sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC025712; AAH25712.1; -.
 DR EMBL; AB082568; BAB92079.1; -.
 DR EMBL; AK074591; BAC11079.1; -.
 DR InterPro; IPR000345; CyC_heme_bind.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 DR PROSITE; PS0028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 266 AA; 29611 MW; 3D15ED256C954E3 CRC64;

Query Match 96.3%; Score 26; DB 4; Length 266;
 Best Local Similarity 83.3%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 DB 109 PALKTK 114

RESULT 5
 QBNN4 PRELIMINARY; PRT; 275 AA.
 AC QBNN4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Striatum;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK081206; BAC38164.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 275 AA; 30822 MW; C25DC8CED9D8204 CRC64;

Query Match 96.3%; Score 26; DB 11; Length 275;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 DB 33 PSLKTK 38

RESULT 6
 O9R7C6 PRELIMINARY; PRT; 318 AA.
 AC O9R7C6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Cell-cycle protein FtsZ (Cell division protein ftsZ)
 DE (Fragment).
 DE FTSZ
 GN Wolbachia pipitensis.
 OS Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Wolbachieae; Wolbachia.
 OX NCBI_TaxID=955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Type I;
 RX MEDLINE=97260975; PubMed=9107051;
 RA Schilthuisen M., Stouthamer R.;
 RT "Horizontal transmission of parthenogenesis-inducing microbes in
 RT Trichogramma wasps."
 RL Proc. R. Soc. Lond., B, Biol. Sci. 264:361-366 (1997).
 CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.
 CC -!- IT SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
 CC THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
 CC AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
 CC BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
 CC -!- SUBUNIT: AGGREGATES TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC ASSEMBLES AT THE INNER SURFACE
 CC OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
 DR EMBL; U74479; AAC64387.1; -.
 DR HSSP; Q57816; 1FSZ.
 DR InterPro; IPR000158; FtsZ.
 DR InterPro; IPR003008; Tubulin_FtsZ.
 DR Pfam; PF00091; tubulin; 1.
 DR Pfam; PF03953; tubulin_C; 1.
 DR PRINTS; PR00423; CELLDIVISFTSZ.
 DR TIGREMS; TIGR00065; ftsz; 1.
 KW Cell cycle; Cell division; GTP-binding; Septation.
 FT NON_TER 1
 FT NON_TER 318 318
 SQ SEQUENCE 318 AA; 34115 MW; C5DA7A10FEC7B216 CRC64;

Query Match 96.3%; Score 26; DB 2; Length 318;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
 DB 306 PTLKTK 311

RESULT 7
 O8F7X4 PRELIMINARY; PRT; 335 AA.
 AC O8F7X4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN LA0818.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.

Query Match 96.3%; Score 26; DB 11; Length 275;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 33 PSLKTK 38

RESULT 10
Q8C569 PRELIMINARY; PRT; 351 AA.

ID Q8C569; AC Q8C569; DT 01-MAR-2003 (TREMBlrel. 23, Created); DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update); DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update); DE Hypothetical protein.; OS Mus musculus (Mouse); OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.; OX NCBI_TaxID=10090; RN [1]; RP SEQUENCE FROM N.A.; RC STRAIN=C57BL/6J; TISSUE=Bone; RX MEDLINE=22354683; PubMed=12466851; RA the FANTOM Consortium; RA the RIKEN Genome Exploration Research Group Phase I & II Team; RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; RL Nature 420:563-573 (2002); DR EMBL; AK079397; BAC37632.1; KW Hypothetical protein.; SQ SEQUENCE 351 AA; 39508 MW; 8082AF895A40FE32 CRC64;

Query Match 96.3%; Score 26; DB 11; Length 351;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 33 PSLKTK 38

RESULT 11
Q96VM9 PRELIMINARY; PRT; 446 AA.

ID Q96VM9; AC Q96VM9; DT 01-DEC-2001 (T-EMBlrel. 19, Created); DT 01-DEC-2001 (T-EMBlrel. 19, Last sequence update); DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update); DE Putative transposase.; OS Fusarium oxysporum f. sp. ciceris; OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; OC Hypocerales; Mitosporic Hypocerales; Fusarium.; OX NCBI_TaxID=62683; RN [1]; RP SEQUENCE FROM N.A.; RC STRAIN=8012; TRANSPOSON=Potci; RA Horman S.R., Bainbridge B.W.; RT "Potci, a hAT family transposable element in Fusarium oxysporum f. sp. ciceris."; RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases. DR EMBL; AY039810; AAK82929.1; DR InterPro; IPR004875; CENP-B. DR InterPro; IPR006600; CENP-B. DR Pfam; PF03184; DDE; 1. DR SMART; SMO0674; CENP; 1. SQ SEQUENCE 446 AA; 50489 MW; B5F1862F7F01ED8A CRC64;

Query Match 96.3%; Score 26; DB 3; Length 446;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |

RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
Ren S.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE011267; AAN49017.1; -
DR DR
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 335 AA; 38329 MW; 56C7119DE62B8DC9 CRC64;

Query Match 96.3%; Score 26; DB 16; Length 335;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 320 PTLKTK 325

RESULT 8
Q9H7B0 PRELIMINARY; PRT; 351 AA.

AC Q9H7B0; DT 01-MAR-2001 (TREMBlrel. 16, Created); DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update); DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update); DE Hypothetical protein FLJ21103.; OS Homo sapiens (Human); OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.; OX NCBI_TaxID=9606; RN [1]; RP SEQUENCE FROM N.A.; RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T., RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y., RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., RA Isegai T., Sugano S.; RT "NEO human cDNA sequencing project."; RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases. DR EMBL; AK024756; BAB14987.1; KW Hypothetical protein.; SQ SEQUENCE 351 AA; 39441 MW; 405A2A9DE03B0308 CRC64;

Query Match 96.3%; Score 26; DB 4; Length 351;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
| | | | |
Db 33 PSLKTK 38

RESULT 9
Q9BPY3 PRELIMINARY; PRT; 351 AA.

AC Q9BPY3; DT 01-JUN-2001 (TREMBlrel. 17, Created); DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update); DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update); DE Similar to hypothetical protein FLJ20635.; OS Homo sapiens (Human); OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.; OX NCBI_TaxID=9606; RN [1]; RP SEQUENCE FROM N.A.; RC TISSUE=Lung, and Cervix; RA Straube B.; RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases. DR EMBL; BC001647; AAH01647.1; -
DR EMBL; BC001340; AAH01340.1; -
SQ SEQUENCE 351 AA; 39499 MW; D298067C1D17E339 CRC64;

Query Match 96.3%; Score 26; DB 4; Length 351;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;

```

Db      113 PALKTK 118
RESULT 12
Q00832      PRELIMINARY;      PRT;      542 AA.
AC Q00832;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Oxyспорum Fot 1 transposon.
OS Fusarium oxysporum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; mitosporic Hypocreales; Fusarium.
OX NCBI_TaxID=5507;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=FOT1-37;
RX MEDLINE=92204124; PubMed=1313143;
RA Daboussi M.J., Langan T., Brygoo Y.;
RT "Fot1, a new family of fungal transposable elements.";
RL Mol. Gen. Genet. 232:12-16(1992).
DR EMBL; X64799; CAA46025.1; -.
DR InterPro; IPR004875; CENP-B.
DR InterPro; IPR006600; CENPB.
DR Pfam; PF03184; DDE; 1.
DR SMART; SM00674; CENPB; 1.
SQ SEQUENCE 542 AA; 60825 MW; 4BC708D1CC858833 CRC64;

Query Match      96.3%; Score 26; DB 3; Length 542;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PXLKTK 6
Db      113 PALKTK 118

RESULT 13
Q9PYX6      PRELIMINARY;      PRT;      568 AA.
AC Q9PYX6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF67.
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum
OS granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=51677;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99434230; PubMed=10502508;
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05181.1; -.
SQ SEQUENCE 568 AA; 66453 MW; 4021A83343407A42 CRC64;

Query Match      96.3%; Score 26; DB 12; Length 568;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PXLKTK 6
Db      560 PSLKTK 565

RESULT 14
Q96A20      PRELIMINARY;      PRT;      577 AA.
AC Q96A20;
DT 01-DEC-2001 (TReMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Middle-chain acyl-CoA synthetase1 (Medium-chain acyl-CoA
DE synthetase).
GN MACS1 OR MACS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2143789; PubMed=11470804;
FU Fujino T., Takei Y.A., Sone H., Ioka R.X., Kamataki A., Magoori K.,
TA Takahashi S., Sakai J., Yamamoto T.T.;
RT "Molecular Identification and Characterization of Two Medium-chain
RT Acyl-CoA Synthetases, MACS1 and the Sa Gene Product.";
RL J. Biol. Chem. 276:35961-35966(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Fujino T., Yamamoto T.T.;
RT "Molecular characterization of medium-chain acyl-CoA synthetase.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB062503; BAB68363.1; JOINED.
DR EMBL; AB062491; BAB68363.1; JOINED.
DR EMBL; AB062492; BAB68363.1; JOINED.
DR EMBL; AB062493; BAB68363.1; JOINED.
DR EMBL; AB062494; BAB68363.1; JOINED.
DR EMBL; AB062495; BAB68363.1; JOINED.
DR EMBL; AB062496; BAB68363.1; JOINED.
DR EMBL; AB062497; BAB68363.1; JOINED.
DR EMBL; AB062498; BAB68363.1; JOINED.
DR EMBL; AB062499; BAB68363.1; JOINED.
DR EMBL; AB062500; BAB68363.1; JOINED.
DR EMBL; AB062501; BAB68363.1; JOINED.
DR EMBL; AB062502; BAB68363.1; JOINED.
DR EMBL; AB059429; BAB64535.1; -.
DR Genew; HGNC:18049; BUCS1.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
SQ SEQUENCE 577 AA; 65259 MW; 380BF4B8911B5B34 CRC64;

Query Match      96.3%; Score 26; DB 4; Length 577;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 PXLKTK 6
Db      180 PSLKTK 185

RESULT 15
Q9BEA2      PRELIMINARY;      PRT;      577 AA.
AC Q9BEA2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Lipocate-activating enzyme precursor.
GN LAE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21369925; PubMed=11382754;
RA Fujiwara K., Takeuchi S., Okamura-Ikeda K., Motokawa Y.;
RT "Purification, Characterization, and cDNA Cloning of Lipocate-
RT activating Enzyme from Bovine Liver.";
RL J. Biol. Chem. 276:28819-28823(2001).
DR EMBL; AB048289; BAB40420.1; -.

```

Tue Aug 12 09:51:00 2003

DR HSSP: P08659; ILCI.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
KW Mitochondrion; Transit peptide.
FT TRANSIT 1 31 MITOCHONDRION.
SQ SEQUENCE 577 AA; 64892 MW; 2768D91B54CBBE CRC64;

Query Match 96.3%; Score 26; DB 6; Length 577;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PXLKTK 6
| | | | |
Db 180 PSLKTK 185

Search completed: August 4, 2003, 16:07:14
Job time : 98 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 16:03:33 ; Search time 30 seconds
 (without alignments)
 8.462 Million cell updates/sec

Title: US-09-492-764B-20
 Perfect score: 27
 Sequence: 1 PXLTKK 6

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

Issued Patents_AA:*
 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
 3: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
 4: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep:*
 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

*** SUMMARIES ***

Result No.	Score	Query Match	Length	DB ID	Description
1	26	96.3	718	1	US-08-190-802A-44
2	26	96.3	718	3	US-08-477-346-44
3	26	96.3	718	4	US-08-473-089-44
4	26	96.3	718	4	US-08-487-072A-44
5	26	96.3	979	3	US-08-878-474-5
6	25	92.6	336	4	US-09-107-532A-4289
7	25	92.6	393	4	US-09-387-418A-10
8	25	92.6	485	2	US-08-477-451-22
9	25	92.6	515	3	US-08-942-012B-24
10	25	92.6	770	1	US-08-369-796-12
11	25	92.6	770	1	US-08-416-581B-1
12	25	92.6	770	1	US-08-416-581B-5
13	25	92.6	770	1	US-08-416-581B-9
14	25	92.6	770	2	US-08-852-091-12
15	25	92.6	770	2	US-08-820-754-12
16	25	92.6	770	3	US-08-956-652-12
17	25	92.6	770	3	US-08-956-869-12
18	25	92.6	770	3	US-09-012-710-8
19	25	92.6	770	3	US-08-948-547-12
20	25	92.6	770	3	US-09-087-465-6
21	25	92.6	770	3	US-09-364-970-3
22	25	92.6	770	3	US-09-164-970-5
23	25	92.6	770	4	US-09-556-273-8
24	25	92.6	770	4	US-08-956-653A-12
25	25	92.6	770	4	US-09-526-542-2
26	25	92.6	770	4	US-09-972-800A-6
27	25	92.6	770	5	PCT-US95-17025-12

ALIGNMENTS

RESULT 1
 US-08-190-802A-44
 ; Sequence 44, Application US/08190802A
 ; Patent No. 5519003
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron, Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: P.O. Box 60850
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306-0850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/190,802A
 ; FILING DATE: 01-FEB-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fabian, Gary R.
 ; REGISTRATION NUMBER: 33,875
 ; REFERENCE/DOCKET NUMBER: 8600-0139
 ; TELEPHONE: (415) 324-0880
 ; TELEFAX: (415) 324-0960
 ; INFORMATION FOR SEQ ID NO: 44:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 718 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: GROUCHO PROTEIN DROSOPH, Fig. 27
 ; US-08-190-802A-44

Sequence 14, Appl
 Sequence 14, Appl
 Sequence 86, Appl
 Sequence 61, Appl
 Sequence 4051, Ap
 Sequence 16, Appl
 Sequence 255, App
 Sequence 255, App
 Sequence 255, App
 Sequence 970, App
 Sequence 973, App
 Sequence 31143, A
 Sequence 414, App
 Sequence 1, Appli
 Sequence 3409, Ap
 Sequence 7785, Ap
 Sequence 24, Appl

Query Match 96.3%; Score 26; DB 1; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLTKK 6

STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/08/473,089
 APPLICATION NUMBER: US/08/473,089
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 44:
 LENGTH: 718 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: GROUCHO PROTEIN DROSOPH, Fig. 27
 US-08-473-089-44

Query Match 96.3%; Score 26; DB 4; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 DB 296 PSLKTK 301

RESULT 4
 US-08-487-072A-44
 Sequence 44, Application US/08487072A
 Patent No. 6423684
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 44:
 LENGTH: 718 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: GROUCHO PROTEIN DROSOPH, Fig. 27
 US-08-477-346-44

Query Match 96.3%; Score 26; DB 3; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 DB 296 PSLKTK 301

RESULT 3
 US-08-473-089-44
 Sequence 44, Application US/08473089
 Patent No. 6342368
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,072A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:

Query Match 96.3%; Score 26; DB 4; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
 DB 296 PSLKTK 301

RESULT 3
 US-08-473-089-44
 Sequence 44, Application US/08473089
 Patent No. 6342368
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,072A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 44:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 718 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: GROUCHO PROTEIN DROSOPH, Fig. 27
 US-08-487-072A-44

Query Match 96.3%; Score 26; DB 4; Length 718;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 296 PSLKTK 301

RESULT 5
 US-08-878-474-5
 ; Sequence 5, Application US/08878474
 ; Patent No. 6133232
 ; GENERAL INFORMATION:
 ; APPLICANT: De Robertis, Edward M.
 ; APPLICANT: Bouwmeester, Tewis
 ; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 ; TITLE OF INVENTION: Factors
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Majestic, Parsons, Siebert & Hsue
 ; STREET: Four Embarcadero Center, Suite 1100
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 94111-4106
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/878,474
 ; FILING DATE: 18-JUN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION NUMBER: US 60/020,150
 ; FILING DATE: 20-JUN-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Siebert, J. Suzanne
 ; REGISTRATION NUMBER: 28,758
 ; REFERENCE/DOCKET NUMBER: 3100.002U51
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/248-5500

Query Match 96.3%; Score 26; DB 3; Length 979;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 296 PSLKTK 301

Db 435 PSLKTK 440
 RESULT 6
 US-09-107-532A-4289
 ; Sequence 4289, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ariniello, Pamela Deneke
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 4289:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 336 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...336
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4289:

Query Match 92.6%; Score 25; DB 4; Length 336;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 93 PXLKTK 98

RESULT 7
 US-09-387-418A-10
 ; Sequence 10, Application US/09387418A
 ; Patent No. 6391572
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Xiaokui
 ; APPLICANT: Wrzeszczynska, Melissa H
 ; APPLICANT: Horvath, Curt M
 ; APPLICANT: Darnell Jr., James E
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR

Query Match 92.6%; Score 25; DB 4; Length 336;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 93 PXLKTK 98

RESULT 7
 US-09-387-418A-10
 ; Sequence 10, Application US/09387418A
 ; Patent No. 6391572
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Xiaokui
 ; APPLICANT: Wrzeszczynska, Melissa H
 ; APPLICANT: Horvath, Curt M
 ; APPLICANT: Darnell Jr., James E
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR

Tue Aug 12 09:50:58 .2003

```

; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-387-418A-10

```

```

Query Match 92.6%; Score 25; DB 4; Length 393;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 PXLKTK 6
Db 327 PYLTKK 332

```

```

RESULT 8
US-08-477-451-22 Application US/08477451
; Sequence 22, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 485 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-451-22

```

```

Query Match 92.6%; Score 25; DB 2; Length 485;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 PXLKTK 6
Db 93 PYLTKK 98

```

```

RESULT 9
US-08-942-012B-24 Application US/08942012B
; Sequence 24, Application US/08942012B

```

```

; Patent No. 6235278
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Lu, Albert
; APPLICANT: Dierks, Peter
; TITLE OF INVENTION: Biological Insect Control Agents Expressing
; TITLE OF INVENTION: Insect-Specific Toxin Genes, Methods and Compositions
; FILE REFERENCE: 28-96a
; CURRENT APPLICATION NUMBER: US/08/942,012B
; CURRENT FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: 08/729,606
; PRIOR FILING DATE: 2000-10-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Helicoverpa zea nuclear polyhedrosis virus
; US-08-942-012B-24

```

```

Query Match 92.6%; Score 25; DB 3; Length 515;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 PXLKTK 6
Db 473 PILKTK 478

```

```

RESULT 10
US-08-369-796-12 Application US/08369796
; Sequence 12, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-369-796-12

```

Query Match 92.6%; Score 25; DB 1; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 704 PYLTKK 709

RESULT 11
 US-08-416-581B-1
 ; Sequence 1, Application US/08416581B
 ; Patent No. 5719042
 ; GENERAL INFORMATION:
 ; APPLICANT: Kishimoto, Tadimitsu
 ; APPLICANT: Akira, Shizuo
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 ; STREET: 2100 Pennsylvania Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,581B
 ; FILING DATE: 04-APR-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-65825/1994
 ; FILING DATE: 04-APR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nakamura, Dean H.
 ; REGISTRATION NUMBER: 33,981
 ; REFERENCE/DOCKET NUMBER: Q-37891
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)293-7060
 ; TELEFAX: (202)293-7860
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 770 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-416-581B-1

Query Match 92.6%; Score 25; DB 1; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 704 PYLTKK 709

RESULT 12
 US-08-416-581B-5
 ; Sequence 5, Application US/08416581B
 ; Patent No. 5719042
 ; GENERAL INFORMATION:
 ; APPLICANT: Kishimoto, Tadimitsu
 ; APPLICANT: Akira, Shizuo
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/416,581B
 FILING DATE: 04-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-65825/1994
 FILING DATE: 04-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Nakamura, Dean H.
 REGISTRATION NUMBER: 33,981
 REFERENCE/DOCKET NUMBER: Q-37891
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-7860
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-581B-5

Query Match 92.6%; Score 25; DB 1; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PXLKTK 6
 Db 704 PYLTKK 709

RESULT 13
 US-08-416-581B-9
 ; Sequence 9, Application US/08416581B
 ; Patent No. 5719042
 ; GENERAL INFORMATION:
 ; APPLICANT: Kishimoto, Tadimitsu
 ; APPLICANT: Akira, Shizuo
 ; TITLE OF INVENTION: TRANSCRIPTION FACTOR APRF
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
 ; STREET: 2100 Pennsylvania Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/416,581B
 ; FILING DATE: 04-APR-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 6-65825/1994
 ; FILING DATE: 04-APR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nakamura, Dean H.
 ; REGISTRATION NUMBER: 33,981

REFERENCE/DOCKET NUMBER: Q-37891
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)293-7060
 TELEFAX: (202)293-7860
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-416-581B-9

Query Match 92.6%; Score 25; DB 1; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 704 PYLKTK 709

RESULT 14
 US-08-852-091-12
 ; Sequence 12, Application US/08852091
 ; Patent No. 5883228
 ; GENERAL INFORMATION:
 ; APPLICANT: James E. Darnell, Jr.
 ; APPLICANT: Zilong Wen
 ; APPLICANT: Curt M. Horvath
 ; APPLICANT: Zhong Zhong
 ; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
 ; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/852,091
 FILING DATE: 06-MAY-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/369,796
 FILING DATE: 06-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-116
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-852-091-12

Query Match 92.6%; Score 25; DB 2; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6

Db 704 PYLKTK 709

RESULT 15
 US-08-820-754-12
 ; Sequence 12, Application US/08820754
 ; Patent No. 5976835
 ; GENERAL INFORMATION:
 ; APPLICANT: Darnell Jr., James E.
 ; APPLICANT: Schindler, Christian W.
 ; APPLICANT: Fu, Xian-yuan
 ; APPLICANT: Wen, Zilong
 ; APPLICANT: Zhong, Zhong
 ; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
 ; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klauber & Jackson
 ; STREET: 411 Hackensack Avenue
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/820,754
 FILING DATE: 19-MAR-1997
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/212,185
 FILING DATE: 11-MAR-1994
 APPLICATION NUMBER: US 07/980,498
 FILING DATE: 23-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/854,296
 FILING DATE: 19-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO US93/02569
 FILING DATE: 19-MAR-1993
 APPLICATION NUMBER: US 08/126,588
 FILING DATE: 24-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-073 CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 770 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-820-754-12

Query Match 92.6%; Score 25; DB 2; Length 770;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 704 PYLKTK 709

Search completed: August 4, 2003, 16:08:36

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2003, 16:05:08 ; Search time 50 Seconds
(without alignments)
14.251 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLKTK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	96.3	54	9	US-09-864-761-44286
2	26	96.3	266	15	US-10-153-668-29
3	26	96.3	350	9	US-09-867-550-1876
4	26	96.3	351	9	US-09-799-777-52
5	26	96.3	896	9	US-09-903-180B-5
6	26	96.3	896	10	US-09-903-171A-5
7	26	96.3	896	10	US-09-903-188A-5
8	26	96.3	896	10	US-09-903-323A-5
9	26	96.3	896	10	US-09-903-325A-5
10	26	96.3	896	10	US-09-903-170C-5
11	26	96.3	979	9	US-09-903-187A-5
12	25	92.6	118	10	US-09-764-864-1349
13	25	92.6	185	15	US-10-156-761-12472
14	25	92.6	312	10	US-09-925-637-48
15	25	92.6	312	15	US-10-084-205-48

16	25	92.6	385	10	US-09-934-332-2
17	25	92.6	393	14	US-10-090-185-10
18	25	92.6	403	15	US-10-156-761-13824
19	25	92.6	428	15	US-10-157-031-139
20	25	92.6	570	10	US-09-877-633-13
21	25	92.6	608	10	US-09-877-633-14
22	25	92.6	608	10	US-09-924-358-8
23	25	92.6	770	15	US-10-045-793-8
24	25	92.6	770	15	US-10-038-010-56
25	25	92.6	793	9	US-09-925-302-780
26	25	92.6	888	15	US-10-156-761-11315
27	25	92.6	1024	15	US-10-211-962-86
28	25	92.6	2120	14	US-10-051-311A-2
29	24	88.9	255	10	US-09-738-626-3575
30	24	88.9	2183	15	US-10-223-070-2
31	23	85.2	10	15	US-10-147-910-21
32	23	85.2	64	9	US-09-815-242-5120
33	23	85.2	81	15	US-10-106-698-7833
34	23	85.2	120	15	US-10-011-931-3
35	23	85.2	205	11	US-09-284-320-1
36	23	85.2	246	11	US-09-880-748-1847
37	23	85.2	246	9	US-09-159-469-24
38	23	85.2	376	9	US-09-798-042-24
39	23	85.2	443	15	US-10-156-761-13902
40	23	85.2	546	15	US-10-043-344-99
41	23	85.2	559	10	US-09-877-633-12
42	23	85.2	597	10	US-09-815-923-14
43	23	85.2	641	14	US-10-108-605-61
44	23	85.2	712	9	US-09-798-042-92
45	23	85.2	718	15	US-10-142-143-20

ALIGNMENTS

RESULT 1

US-09-864-761-44286
; Sequence 44286, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemlica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

Best Local Similarity 83.3%; Pred. No. 1.2e+02; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
Db 109 PALKTK 114

RESULT 3

US-09-867-550-1876
; Sequence 1876, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells an
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1876
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1876

Query Match 96.3%; Score 26; DB 9; Length 350;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
Db 33 PSLKTK 38

RESULT 4

US-09-799-777-52
; Sequence 52, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Purvi
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

Query Match 96.3%; Score 26; DB 9; Length 54;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
Db 34 PSLKTK 39

RESULT 2

US-10-153-668-29
; Sequence 29, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STRA6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-29

Query Match 96.3%; Score 26; DB 15; Length 266;

APPLICATION NUMBER: US/09/002,485
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BILLINGS, LUCY J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0459 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 52:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 351 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: ENDANOT01
 CLONE: 2452208
 SEQUENCE DESCRIPTION: SEQ ID NO: 52 :

US-09-799-777-52
 Query Match 96.3%; Score 26; DB 9; Length 351;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 33 PSLKTK 38

RESULT 5
 US-09-903-180B-5
 ; Sequence 5, Application US/09903180B
 ; Patent No. US20020099171A1
 ; GENERAL INFORMATION:
 ; APPLICANT: De Robertis, Edward M.
 ; APPLICANT: Boumeester, Tewis
 ; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 ; TITLE OF INVENTION: Factors
 ; FILE REFERENCE: 510015-256
 ; CURRENT APPLICATION NUMBER: US/09/903,180B
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 60/020,150
 ; PRIOR FILING DATE: 1996-06-20
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 896
 ; TYPE: PRT
 ; ORGANISM: Xenopus
 US-09-903-180B-5

Query Match 96.3%; Score 26; DB 9; Length 896;
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 435 PSLKTK 440

RESULT 6
 US-09-903-171A-5
 ; Sequence 5, Application US/09903171A
 ; Patent No. US20020123613A1
 ; GENERAL INFORMATION:
 ; APPLICANT: De Robertis, Edward M.
 ; APPLICANT: Boumeester, Tewis
 ; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 ; TITLE OF INVENTION: Factors
 ; FILE REFERENCE: 510015-260
 ; CURRENT APPLICATION NUMBER: US/09/903,171A
 ; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: US 60/020,150
 ; PRIOR FILING DATE: 1996-06-20
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 896
 ; TYPE: PRT
 ; ORGANISM: Xenopus
 US-09-903-171A-5

Query Match 96.3%; Score 26; DB 10; Length 896;
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 435 PSLKTK 440

RESULT 7
 US-09-903-188A-5
 ; Sequence 5, Application US/09903188A
 ; Patent No. US20020128439A1
 ; GENERAL INFORMATION:
 ; APPLICANT: De Robertis, Edward M.
 ; APPLICANT: Boumeester, Tewis
 ; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 ; TITLE OF INVENTION: Factors
 ; FILE REFERENCE: 510015-258
 ; CURRENT APPLICATION NUMBER: US/09/903,188A
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 60/020,150
 ; PRIOR FILING DATE: 1996-06-20
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 896
 ; TYPE: PRT
 ; ORGANISM: Xenopus
 US-09-903-188A-5

Query Match 96.3%; Score 26; DB 10; Length 896;
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 435 PSLKTK 440

RESULT 8
 US-09-903-323A-5
 ; Sequence 5, Application US/09903323A
 ; Patent No. US20020128440A1
 ; GENERAL INFORMATION:
 ; APPLICANT: De Robertis, Edward M.
 ; APPLICANT: Boumeester, Tewis
 ; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 ; TITLE OF INVENTION: Factors
 ; FILE REFERENCE: 510015-261
 ; CURRENT APPLICATION NUMBER: US/09/903,323A
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: US 60/020,150
 ; PRIOR FILING DATE: 1996-06-20
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 896
 ; TYPE: PRT
 ; ORGANISM: Xenopus
 US-09-903-323A-5

Query Match 96.3%; Score 26; DB 10; Length 896;

Best Local Similarity 83.3%; Pred. No. 3.8e+02; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
Db 435 PSLKTK 440

RESULT 9
US-09-903-325A-5
; Sequence 5, Application US/09903325A
; Patent No. US20020128441A1
; GENERAL INFORMATION:
; APPLICANT: Bouwmeester, Edward M.
; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
; TITLE OF INVENTION: Factors
; FILE REFERENCE: 510015-257
; CURRENT APPLICATION NUMBER: US/09/903,325A
; CURRENT FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/020,150
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Xenopus
US-09-903-325A-5

Query Match 96.3%; Score 26; DB 10; Length 896;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
Db 435 PSLKTK 440

RESULT 10
US-09-903-170C-5
; Sequence 5, Application US/09903170C
; Patent No. US20020156249A1
; GENERAL INFORMATION:
; APPLICANT: De Robertis, Edward M.
; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
; TITLE OF INVENTION: Factors
; FILE REFERENCE: 510015-259
; CURRENT APPLICATION NUMBER: US/09/903,170C
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 60/020,150
; PRIOR FILING DATE: 1996-06-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Xenopus
US-09-903-170C-5

Query Match 96.3%; Score 26; DB 10; Length 896;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
Db 435 PSLKTK 440

RESULT 11
US-09-903-187A-5
; Sequence 5, Application US/09903187A

; Patent No. US20020099172A1
; GENERAL INFORMATION:
; APPLICANT: De Robertis, Edward M.
; APPLICANT: Bouwmeester, Tewis
; TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
; TITLE OF INVENTION: Factors
; FILE REFERENCE: 510015-248
; CURRENT APPLICATION NUMBER: US/09/903,187A
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/552,988
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: US 08/878,474
; PRIOR FILING DATE: 1997-06-19
; PRIOR APPLICATION NUMBER: US 60/020,150
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 979
; TYPE: PRT
; ORGANISM: Xenopus
US-09-903-187A-5

Query Match 96.3%; Score 26; DB 9; Length 979;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
Db 435 PSLKTK 440

RESULT 12
US-09-764-864-1349
; Sequence 1349, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1349
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1349

Query Match 92.6%; Score 25; DB 10; Length 118;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PXLKTK 6
Db 74 PLLKTK 79

RESULT 13
US-10-156-761-12472
; Sequence 12472, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 12472
 ; LENGTH: 185
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-12472

Query Match 92.6%; Score 25; DB 15; Length 185;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 7 PRLKTK 12

RESULT 14
 US-09-925-637-48
 ; Sequence 48, Application US/09925637
 ; Patent No. US20020103338A1
 ; GENERAL INFORMATION:

; APPLICANT: Choi
 ; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
 ; FILE REFERENCE: PB560
 ; CURRENT APPLICATION NUMBER: US/09/925,637
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/23773
 ; PRIOR FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: US 60/151,933
 ; PRIOR FILING DATE: 1999-09-01
 ; PRIOR APPLICATION NUMBER: US 08/781,986
 ; PRIOR FILING DATE: 1997-01-03
 ; PRIOR APPLICATION NUMBER: US 08/956,171
 ; PRIOR FILING DATE: 1997-10-20
 ; PRIOR APPLICATION NUMBER: US 60/009,861
 ; PRIOR FILING DATE: 1996-01-06
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 48
 ; LENGTH: 312
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-925-637-48

Query Match 92.6%; Score 25; DB 10; Length 312;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 7 PRLKTK 12

RESULT 15
 US-10-084-205-48
 ; Sequence 48, Application US/10084205
 ; Publication No. US20030049648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi, Gil

; TITLE OF INVENTION: 37 Staphylococcus aureus Genes and Polypeptides
 ; FILE REFERENCE: PB515P1
 ; CURRENT APPLICATION NUMBER: US/10/084,205
 ; CURRENT FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: PCT/US00/23773
 ; PRIOR FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 60/151,933
 ; PRIOR FILING DATE: 1999-09-01
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 48
 ; LENGTH: 312
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-10-084-205-48

Query Match 92.6%; Score 25; DB 15; Length 312;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PXLKTK 6
 Db 7 PRLKTK 12

Search completed: August 4, 2003, 16:09:33
 Job time : 51 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 9, 2003, 18:20:23 ; Search time 1484 Seconds
(without alignments)
165.403 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLTKK 6

Scoring table: BLOSUM62
Ygapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2/1/USFO_spool/US09492764/runat_04082003_142326_11126/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09492764 @CGN 1 2372 @runat_04082003_142326_11126 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	26	96.3	115	8	ATH523373 Arabidops
2	26	96.3	115	8	ATH532053 Arabidops
3	26	96.3	152	11	HSP644H11 H.sapiens
4	26	96.3	183	6	A74872 Sequence 55
5	26	96.3	183	6	A77851 Sequence 55
6	26	96.3	208	6	BD042667 Sequence 55
7	26	96.3	213	11	G05278 human STS W
8	26	96.3	222	6	AX523293 Sequence
9	26	96.3	224	8	AY023300 Oryza sat
10	26	96.3	243	6	AR240305 Sequence
11	26	96.3	244	9	AB06249S03 Homo sapi
12	26	96.3	255	6	AX210836 Sequence
13	26	96.3	266	6	BD049717 Sequence
14	26	96.3	271	8	ATH529154 Arabidops
15	26	96.3	280	4	AF322622 Macropus
16	26	96.3	281	9	HS9E2R H.sapiens C
17	26	96.3	288	6	BD059656 Sacrated
18	26	96.3	290	11	BV050278 S212P6880
19	26	96.3	290	11	G47309 Homo sapi
20	26	96.3	298	9	AF230316 Homo sapi
21	26	96.3	307	6	AX522520 Sequence
22	26	96.3	308	6	AX522339 Sequence
23	26	96.3	313	5	S48571 {fingerprin
24	26	96.3	332	6	AR147974 Sequence
25	26	96.3	332	6	AR219698 Sequence
26	26	96.3	332	6	AX316816 Sequence
27	26	96.3	332	6	BD084349 Compositi
28	26	96.3	344	11	BV062171 S212P6030
29	26	96.3	346	11	HS95G17A H.sapiens
30	26	96.3	350	6	AR051975 Sequence
31	26	96.3	350	6	I52149 Sequence 4B
32	26	96.3	352	11	HS320ZG1 H.sapiens D
33	26	96.3	352	11	HS320ZG1 H.sapiens D
34	26	96.3	356	11	AU027736 Rattus no
35	26	96.3	388	11	DM4B3S D. melanoga
36	26	96.3	392	10	S80586 igfbp-4:ins
37	26	96.3	397	6	AX072216 Sequence
38	26	96.3	404	6	BD038282 Sequence
39	26	96.3	404	6	BD046113 Sequence
40	26	96.3	410	6	AX193113 Sequence
41	26	96.3	410	11	G07537 human STS S
42	26	96.3	416	6	AX334279 Sequence
43	26	96.3	421	11	G57191 SHCG-103006
44	26	96.3	424	11	G41717 Z11410 Zebri
45	26	96.3	428	1	UBA518264 Unidentif

ALIGNMENTS

RESULT 1

ATH523373 115 bp DNA linear PLN 29-MAR-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 DEFINITION 057F11.
 ACCESSION AJ523373
 VERSION left border; T-DNA flanking sequence.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1
 AUTHORS Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
 Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
 Lepiniec,L., Caboche,M. and Lecharny,A.
 T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 115)
 AUTHORS Balzergue,S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
 http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (http://www.genoplante.com and
 http://genoplante-info.infobiogen.fr).

FEATURES
 source
 1..115
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultiivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="250D07"
 /clone.lib="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature
 1..115
 /note="T-DNA flanking sequence
 left border" 13 g 42 t

BASE COUNT 49 a 11 c 13 g 42 t
 ORIGIN

Alignment Scores: Length: 115
 Pred. No.: 224 Matches: 5
 Score: 26.00 Conservative: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 83.33% Indels: 0
 Query Match: 96.30% Gaps: 0
 DB: 8

US-09-492-764B-20 (1-6) x ATH532053 (1-115)

QY 1 PRO***LeuLy5ThrLy5 6
 ||| |||||
 Db 8 CCAACTCTAAAAACTAAA 25

RESULT 3
 HSP644H11 152 bp DNA linear STS 21-MAY-1998
 LOCUS H.sapiens flow-sorted chromosome 1 HindIII fragment, SC1PE44H11,
 DEFINITION sequence tagged site.
 ACCESSION AL010011
 VERSION AL010011.1 GI:2665173
 KEYWORDS STS; single read.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 152)
 AUTHORS Gregory,S., Kettleborough,R., Langford,C., Ross,M.T. and Hunt,S.E.
 TITLE Direct Submission

ATH523373 115 bp DNA linear PLN 29-MAR-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 DEFINITION 057F11.
 ACCESSION AJ523373
 VERSION left border; T-DNA flanking sequence.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1
 AUTHORS Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
 Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
 Lepiniec,L., Caboche,M. and Lecharny,A.
 T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 115)
 AUTHORS Balzergue,S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
 http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (http://www.genoplante.com and
 http://genoplante-info.infobiogen.fr).

FEATURES
 source
 1..115
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultiivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="057F11"
 /clone.lib="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature
 1..115
 /note="T-DNA flanking sequence
 left border" 13 g 42 t

BASE COUNT 49 a 11 c 13 g 42 t
 ORIGIN

Alignment Scores: Length: 115
 Pred. No.: 224 Matches: 5
 Score: 26.00 Conservative: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 83.33% Indels: 0
 Query Match: 96.30% Gaps: 0
 DB: 8

US-09-492-764B-20 (1-6) x ATH523373 (1-115)

QY 1 PRO***LeuLy5ThrLy5 6
 ||| |||||
 Db 8 CCAACTCTAAAAACTAAA 25

RESULT 2
 ATH532053 115 bp DNA linear PLN 29-MAR-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone
 DEFINITION 250D07.
 ACCESSION AJ532053
 VERSION AJ532053.1 GI:26800313
 KEYWORDS left border; T-DNA flanking sequence.

JOURNAL COMMENT Submitted (03-DEC-1997) E-mail contact: humquery@sanger.ac.uk
 Vector: pBSIISK+ Marker at:BS33113PS (Primer A : ACCTGTGATCATGCCACTGA; Primer B : CTGCTATCCTCAGCTGCG; ampImer: size : 135 bp) was mapped to chromosome 1 using Radiation Hybrid panel Genebridge 4 (GB4).

FEATURES location/Qualifiers source
 1..152 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="SC1PE44H11"
 /sex="Female"
 /tissue type="EBV lymphoblastoid cell line"
 /clone_lib="SC1PE"
 /dev_stages="adult"
 56 a 36 c 31 g 29 t

Alignment Scores: Length: 152
 Score: 290 Matches: 5
 Percent Similarity: 83.33% Conservatative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 11 Gaps: 0

US-09-492-764B-20 (1-6) x HSPE44H11 (1-152)
 QY 1 Pro***LeuLystrLys 6
 ||| |||||
 54 CCATCCCTAAAAACAAA 71

RESULT 4 A74872 183 bp DNA linear PAT 15-OCT-1999
 LOCUS A74872 Sequence 558 from Patent WO9401548.
 DEFINITION A74872
 ACCESSION A74872
 VERSION A74872.1 GI:6064886
 KEYWORDS unidentified
 ORGANISM unclassified
 SOURCE unclassified
 REFERENCE 1 (bases 1 to 183)
 AUTHORS Sibson,D.R. and Gross,J.
 TITLE HUMAN NUCLEIC ACID FRAGMENTS, ISOLATED FROM BRAIN ADRENAL TISSUE,
 PLACENTA OR BONE NARROW
 JOURNAL Patent: WO 9401548-A 558 20-JAN-1994;
 MEDICAL RES COUNCIL (GB); SIBSON DAVID ROSS (GB)

FEATURES location/Qualifiers source
 1..183 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 62 a 27 c 38 g 56 t

Alignment Scores: Length: 183
 Pred. No.: 344 Matches: 5
 Score: 2600 Conservatative: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 83.33% Indels: 0
 Query Match: 96.30% Gaps: 0
 DB: 6

US-09-492-764B-20 (1-6) x A74872 (1-183)
 QY 1 Pro***LeuLystrLys 6
 ||| |||||
 134 CCTTCACTGAAAAACAAA 117

RESULT 5 A77851 183 bp DNA linear PAT 19-OCT-1999
 LOCUS A77851 Sequence 558 from Patent EP0587279.
 DEFINITION A77851
 ACCESSION A77851
 VERSION A77851.1 GI:6089516
 KEYWORDS unclassified
 ORGANISM unclassified
 REFERENCE 1 (bases 1 to 183)
 AUTHORS Sibson,D.R. and Hadfield,K.M.
 TITLE HUMAN NUCLEIC ACID FRAGMENTS ISOLATED FROM BRAIN, ADRENAL TISSUE,
 PLACENTA OR BONE NARROW AND THEIR USE
 JOURNAL Patent: EP 0587279-A 558 16-MAR-1994;
 MEDICAL RES COUNCIL (GB)

FEATURES location/Qualifiers source
 1..183 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 62 a 27 c 38 g 56 t

Alignment Scores: Length: 183
 Pred. No.: 344 Matches: 5
 Score: 2600 Conservatative: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 83.33% Indels: 0
 Query Match: 96.30% Gaps: 0
 DB: 6

US-09-492-764B-20 (1-6) x A77851 (1-183)
 QY 1 Pro***LeuLystrLys 6
 ||| |||||
 134 CCTTCACTGAAAAACAAA 117

RESULT 6 BD042667 208 bp DNA linear PAT 27-AUG-2002
 LOCUS BD042667 Sequence tag and encoded human protein.
 DEFINITION BD042667
 ACCESSION BD042667
 VERSION BD042667.1 GI:22584409
 KEYWORDS JP 2001269182-A/18913.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 208)
 AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
 TITLE Sequence tag and encoded human protein
 JOURNAL Patent: JP 2001269182-A 18913 02-OCT-2001;
 GENSET
 COMMENT OS Homo sapiens (human)
 PN JP 2001269182-A/18913
 PD 02-OCT-2001
 PF 24-FEB-2000 JP 2000118773
 PR 26-FEB-1999 US 60/122487
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES,
 PC C12N15/09.C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
 C12N5/10,
 PC C12P21/02.C12P21/08.C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
 G06F15/40
 CC
 FH Key Location/Qualifiers
 1..208 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 50 a 45 c 30 g 83 t

Alignment Scores:
 Pred. No.: 387 Length: 208
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0

US-09-492-764B-20 (1-6) x BD042667 (1-208)
 QY 1 Pro**LeuLysThrLys 6
 Db 35 CCTAGTTTAAAGACAAG 18

RESULT 7
 G05278/c G05278 213 bp DNA linear STS 19-OCT-1995
 LOCUS human STS WI-5654, sequence tagged site.

DEFINITION G05278
 ACCESSION G05278
 VERSION G05278.1 GI:852194
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 213)
 AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Whitehead Institute/MIT Center for Genome Research; Random Genome
 JOURNAL Wide STS
 REFERENCE 2 (bases 1 to 213)
 AUTHORS Whitehead Institute/MIT Center for Genome Research; Physically
 TITLE Mapped STS
 JOURNAL Unpublished (1995)
 COMMENT

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: CTTAATGAAAACAATGCCAGAGC
 Primer B: TGCAAAATGTGGAATAATCTGG
 STS size: 150
 PCR Profile:
 Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng
 Primer: each 5 pM
 dNTPs: each 4 nM
 Tag Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:
 MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 9.3
 Location/Qualifiers:
 1..213
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="888_C_2; 890_A_8; 910_C_4; 942_A_6; 961_H_9"

FEATURES
 SOURCE
 Location/Qualifiers:
 1..213
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="888_C_2; 890_A_8; 910_C_4; 942_A_6; 961_H_9"

STS primer bind 60..209
 primer_bind 60..82
 complement(188..209)
 BASE COUNT 64 a 39 c 38 g 72 t
 ORIGIN

Alignment Scores:
 Pred. No.: 396 Length: 213
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 11 Gaps: 0

US-09-492-764B-20 (1-6) x G05278 (1-213)
 QY 1 Pro**LeuLysThrLys 6
 Db 113 CCATCTTAAAAACAAAA 96

RESULT 8
 AX523293/c AX523293 222 bp DNA linear PAT 24-OCT-2002
 LOCUS Sequence 963 from Patent WO02064731.
 DEFINITION AX523293
 ACCESSION AX523293
 VERSION AX523293.1 GI:24412247
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL Telexman,A., Anson,R., Tuijnder,M. and Susini,L.
 REFERENCE 2
 AUTHORS Sequences involved in phenomena of tumour suppression, tumour
 TITLE reversion, apoptosis and/or virus resistance and their use as
 JOURNAL medicines
 Patent: WO 02064731-A 963 22-AUG-2002;
 Molecular Engines Laboratories (FR)
 Location/Qualifiers

FEATURES
 source
 1..222
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 61 a 36 c 42 g 83 t
 ORIGIN

Alignment Scores:
 Pred. No.: 411 Length: 222
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0

US-09-492-764B-20 (1-6) x AX523293 (1-222)
 QY 1 Pro**LeuLysThrLys 6
 Db 214 CCATCTTAAAAACTAAA 197

RESULT 9
 AY023300/c AY023300 224 bp DNA linear PLN 07-FEB-2001
 LOCUS Oryza sativa microsatellite MRG5625 containing (GGT)X8, genomic
 DEFINITION sequence.

ACCESSION AY023300
 VERSION AY023300.1 GI:127066516
 KEYWORDS Oryza sativa
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 REFERENCE 1
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 TITLE Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 JOURNAL Ehrhartoideae; Oryzaceae; Oryza.


```

REFERENCE 1 (bases 1 to 224)
AUTHORS Tao, N., Barbazuk, W. B., Liu, J., Wu, K. and Barry, G. F.
TITLE Simple sequence repeats from Monsanto rice genomic sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao, N., Barbazuk, W. B., Liu, J., Wu, K. and Barry, G. F.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh
Blvd., Creve Coeur, MO 63167, USA
COMMENT Derived from rice genomic sequences generated from the Monsanto
Rice Genome Sequencing project. Please see
http://www.rice-research.org for more information. The sequence
data were produced primarily in the laboratories of Dr. Leroy Hood
at the University of Washington in Seattle.
FEATURES
source
1..224
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
repeat_region
1..224
/note="microsatellite MRG5625"
/rpt_type=tandem
/rpt_unit=gggt
BASE COUNT 32 a 57 c 93 g 42 t
ORIGIN
Alignment Scores: Length: 224
Pred. No.: 415 Matches: 5
Score: 26.00 Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 8

US-09-492-764B-20 (1-6) x AY023300 (1-224)
QY 1 Pro***LeuLysThrLys 6
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 CCCTCGCTAAAGACCAA 189

RESULT 10
AR240305 AR240305 243 bp DNA linear PAT 20-DEC-2002
LOCUS Sequence 61 from patent US 6468758.
DEFINITION AR240305
ACCESSION AR240305 GI:27285385
VERSION AR240305.1
KEYWORDS Unknown.
SOURCE Unclassified.
ORGANISM Ununknown.
REFERENCE 1 (bases 1 to 243)
AUTHORS Benson, D. R., Lodes, M. J., Mitcham, J. L. and King, G. E.
TITLE Compositions and methods for ovarian cancer therapy and diagnosis
JOURNAL Patent: US 6468758-A 61 22-OCT-2002;
FEATURES
Location/Qualifiers
source
1..243
/organism="unknown"
BASE COUNT 88 a 42 c 53 g 60 t
ORIGIN
Alignment Scores: Length: 243
Pred. No.: 447 Matches: 5
Score: 26.00 Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 6

US-09-492-764B-20 (1-6) x AR240305 (1-243)
QY 1 Pro***LeuLysThrLys 6
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 CCAGCCTTAAGACCAA 162

```

```

RESULT 11
AB06249S03
LOCUS Homo sapiens MACS1 gene for middle-chain acyl-CoA synthetase1, exon
DEFINITION Homo sapiens MACS1 gene for middle-chain acyl-CoA synthetase1, exon
3.
ACCESSION AB062493 GI:15706410
VERSION AB062493.1
KEYWORDS Homo sapiens
SEGMENT 3 of 13
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Fujino, T., Takei, Y. A., Sone, H., Ioka, R. X., Kamataki, A., Magoori, K.,
Takahashi, S., Sakai, J. and Yamamoto, T. T.
TITLE Molecular identification and characterization of two medium-chain
acyl-CoA synthetases, MACS1 and the Sa gene product
JOURNAL J. Biol. Chem. 276 (38), 35961-35966 (2001)
MEDLINE 21443789
PUBMED 11470804
REFERENCE 2 (bases 1 to 244)
AUTHORS Fujino, T.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2001) Takahiro Fujino, Tohoku University Gene
Research Center; 1-1 Teutsumidori-Amamiya, Sendai 981-8555, Japan.
(E-mail: tfujino@biochem.tohoku.ac.jp, Tel: 81-22-717-8875,
Fax: 81-22-717-8877)
FEATURES
Location/Qualifiers
source
1..244
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
19..226
/feature="MACS1"
/product="middle-chain acyl-CoA synthetase1"
/note="CDS is reported in Acc#:AB062503"
/number=3
/EC_number="6.2.1.2"
BASE COUNT 51 a 69 c 56 g 68 t
ORIGIN
Alignment Scores: Length: 244
Pred. No.: 449 Matches: 5
Score: 26.00 Conservative: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 9

US-09-492-764B-20 (1-6) x AB06249S03 (1-244)
QY 1 Pro***LeuLysThrLys 6
||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 CCCTCTCGAAACCCAAG 170

RESULT 12
AX210836/2
LOCUS AX210836 255 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 478 from Patent WO0157058.
ACCESSION AX210836
VERSION AX210836.1 GI:15425096
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
REFERENCE 1 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
AUTHORS Rosenthal, A., Hinzmann, B., Schaefer, R., Zuber, J., Tchernitsa, O.,
Grips, M., Hellriegel, M., Schmitz, A. C. and Sers, C.

```

TITLE Detection of differential gene expression
JOURNAL Patent: WO 0157058-A 478 09-AUG-2001;
 Metagen Gesellschaft fuer Genomforschung mbH (DE)
FEATURES Location/Qualifiers
 source 1..255
 /organism="Rattus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:10118"
BASE COUNT 67 a 49 c 56 g 81 t 2 others
ORIGIN
 Alignment Scores:
 Pred. No.: 467 Length: 255
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0
 US-09-492-764B-20 (1-6) x AX210836 (1-255)
QY 1 Pro***LeuLysThrLys 6
 ||| |||||
Db 62 CCCAGTCTAAACAAA 45
RESULT 13
LOCUS BD049717/C 266 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD049717
VERSION BD049717.1 GI:22591459
KEYWORDS JP 2001269182-A/25963.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 266)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 25963 02-OCT-2001;
COMMENT GENSPT.
 OS Homo sapiens (human)
 PN JP 2001269182-A/25963
 PD 02-OCT-2001
 PF 24-FEB-2000 JP 2000118773
 PR 26-FEB-1999 US 60/122487
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
 PJ JORDAN
 PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
 C12N5/10,
 PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
 G06F15/40
CC G06F15/40
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
 source 1..266
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
BASE COUNT 73 a 37 c 71 g 84 t 1 others
ORIGIN
 Alignment Scores:
 Pred. No.: 486 Length: 266
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0
 US-09-492-764B-20 (1-6) x BD049717 (1-266)
QY 1 Pro***LeuLysThrLys 6
 ||| |||||
Db 62 CCCAGTCTAAACAAA 45
RESULT 14
LOCUS ATH529154/C 271 bp DNA linear PLN 29-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 179F08.
ACCESSION AJ529154
VERSION AJ529154.1 GI:26797414
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1
AUTHORS Brunaud,V., Balzerque,S., Dubreucq,B., Aubourg,S., Samson,F.,
 Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
 Lepiniec,L., Caboche,M. and Lecharny,A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences
 of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE 2 (bases 1 to 271)
AUTHORS Balzerque,S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
 Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
 plants from INRA (Versailles). The DNA fragment(s) resulting from
 the PCR were directly sequenced from the left or the right border
 to determine the genomic sequence flanking the insertion. T-DNA
 derived sequences were removed. Information to order the
 corresponding mutant line and a link to a database providing a
 graphical display of the insertion site are available at
 http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
 been generated in the framework of the French plant genomics
 program 'Genoplante' (http://www.genoplante.com and
 http://genoplante-info.infobiogen.fr).
FEATURES Location/Qualifiers
 source 1..271
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /culivar="Wassillewskija"
 /db_xref="taxon:3702"
 /clone="I79F08"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
misc_feature 1..271
 /note="T-DNA flanking sequence.
 left border"
BASE COUNT 75 a 37 c 66 g 93 t
ORIGIN
 Alignment Scores:
 Pred. No.: 494 Length: 271
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 8 Gaps: 0
 US-09-492-764B-20 (1-6) x ATH529154 (1-271)
QY 1 Pro***LeuLysThrLys 6
 ||| |||||
Db 265 CCGAGTCTAAACAAA 248
RESULT 15
LOCUS AF322622 280 bp DNA linear MAM 23-JAN-2002
DEFINITION Macropus giganteus microsatellite G12-6 sequence.

TITLE Detection of differential gene expression
JOURNAL Patent: WO 0157058-A 478 09-AUG-2001;
 Metagen Gesellschaft fuer Genomforschung mbH (DE)
FEATURES Location/Qualifiers
 source 1..255
 /organism="Rattus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:10118"
BASE COUNT 67 a 49 c 56 g 81 t 2 others
ORIGIN
 Alignment Scores:
 Pred. No.: 467 Length: 255
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0
 US-09-492-764B-20 (1-6) x AX210836 (1-255)
QY 1 Pro***LeuLysThrLys 6
 ||| |||||
Db 62 CCCAGTCTAAACAAA 45
RESULT 13
LOCUS BD049717/C 266 bp DNA linear PAT 27-AUG-2002
DEFINITION Sequence tag and encoded human protein.
ACCESSION BD049717
VERSION BD049717.1 GI:22591459
KEYWORDS JP 2001269182-A/25963.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 266)
AUTHORS Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
TITLE Sequence tag and encoded human protein
JOURNAL Patent: JP 2001269182-A 25963 02-OCT-2001;
COMMENT GENSPT.
 OS Homo sapiens (human)
 PN JP 2001269182-A/25963
 PD 02-OCT-2001
 PF 24-FEB-2000 JP 2000118773
 PR 26-FEB-1999 US 60/122487
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
 PJ JORDAN
 PC C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC
 C12N5/10,
 PC C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC
 G06F15/40
CC G06F15/40
FH Key Location/Qualifiers.
FEATURES Location/Qualifiers
 source 1..266
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
BASE COUNT 73 a 37 c 71 g 84 t 1 others
ORIGIN
 Alignment Scores:
 Pred. No.: 486 Length: 266
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 6 Gaps: 0
 US-09-492-764B-20 (1-6) x BD049717 (1-266)
QY 1 Pro***LeuLysThrLys 6
 ||| |||||
Db 62 CCCAGTCTAAACAAA 45

ACCESSION AF322622
 VERSION AF322622.1 GI:13785602
 KEYWORDS
 SOURCE Macropus giganteus (eastern gray kangaroo)
 ORGANISM
 Macropus giganteus (eastern gray kangaroo)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 REFERENCE 1 (bases 1 to 280)
 AUTHORS Zenger, K.R. and Cooper, D.W.
 TITLE A set of highly polymorphic microsatellite markers developed for
 the eastern gray kangaroo (Macropus giganteus)
 JOURNAL Mol. Ecol. Notes 1 (1-2), 98-100 (2001)
 REFERENCE 2 (bases 1 to 280)
 AUTHORS Zenger, K.R. and Cooper, D.W.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2000) Biological Sciences, Macquarie University,
 Herring Road, Sydney, NSW 2109, Australia

FEATURES
 source
 1..280
 /organism="Macropus giganteus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9317"
 /clone="G12-6"
 repeat_region
 210..253
 /note="microsatellite G12-6"
 /rpt_type=tandem
 /rpt_unit=ca

BASE COUNT 89 a 79 c 47 g 65 t
 ORIGIN

Alignment Scores:
 Pred. No.: 509 Length: 280
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 4 Gaps: 0

US-09-492-764B-20 (1-6) X AF322622 (1-280)
 QY 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 133 CCCTCACTCAAAACAAAA 150

Search completed: August 9, 2003, 18:45:30
 Job time: 1487 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 7, 2003, 14:26:08 ; Search time 177 Seconds
(without alignments)
91.506 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLTKK 6

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPFO_epool/US09492764/runat_04082003_142326_11116/app_query.fasta_1.199
-DB=N Geneseg 19Jun03 -QMT=fastacp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09492764@cgn_1_0_@runat_04082003_142326_11116 -NCFU=6 -ICFU=3
-NO_MMAPP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseg_19Jun03.*
1: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseg/geneseg-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	26	96.3	51	22	AAL31879	Human SNP oligonuc
C 2	26	96.3	162	22	ABA71554	Human foetal liver
C 3	26	96.3	162	22	AAK19884	Human brain expres
C 4	26	96.3	162	22	AAK45914	Human bone marrow
C 5	26	96.3	162	22	AAI51832	Probe #20518 used
C 6	26	96.3	162	22	ABS45612	Human liver single
C 7	26	96.3	162	24	ABS20199	Human genome-deriv
C 8	26	96.3	183	15	AAO76958	Human genome fragm
C 9	26	96.3	208	21	AAI18922	Human secreted pro
C 10	26	96.3	213	20	AAK88587	Human chromosome 1
C 11	26	96.3	222	25	ABZ79256	Tumour suppression
C 12	26	96.3	222	25	ABZ09803	Human oligonucleot
C 13	26	96.3	227	25	ABZ78118	Human Suppression
C 14	26	96.3	243	24	ABT06899	Human ovarian canc
C 15	26	96.3	243	24	AAZ30827	Clone G3106B cDNA
C 16	26	96.3	243	25	ABX72777	Human ovarian carc
C 17	26	96.3	255	22	AAH81969	Rat differential t
C 18	26	96.3	266	21	AAZ25972	Human secreted pro
C 19	26	96.3	288	20	AAV87538	EST clone DD413
C 20	26	96.3	298	24	ABS67305	Breast specific pol
C 21	26	96.3	305	22	AAI13459	Human breast canc
C 22	26	96.3	306	22	AAI08477	Human breast canc
C 23	26	96.3	307	25	ABZ78483	Tumour suppression
C 24	26	96.3	307	25	ABZ09030	Human oligonucleot
C 25	26	96.3	308	25	ABZ78302	Tumour suppression
C 26	26	96.3	308	25	ABZ08849	Human oligonucleot
C 27	26	96.3	313	22	AAI26373	Human breast canc
C 28	26	96.3	327	22	AAI16908	Human breast canc
C 29	26	96.3	332	18	AAH83342	Breast cancer tumo
C 30	26	96.3	332	19	AAV69010	DNA molecule encod
C 31	26	96.3	332	21	AAK80865	Human breast tumou
C 32	26	96.3	332	24	ABK46755	Human breast tumou
C 33	26	96.3	332	24	AAAS98711	Breast tumour-spec
C 34	26	96.3	351	22	AAI87605	Human polynucleoti
C 35	26	96.3	351	25	ABX39441	Bovine EST associa
C 36	26	96.3	352	24	ABT06935	Human ovarian carc
C 37	26	96.3	352	25	ABX72813	Human ovarian carc
C 38	26	96.3	362	22	AAI85902	Human polynucleoti
C 39	26	96.3	369	16	AAI19142	Human gene signatu
C 40	26	96.3	369	22	AAI90872	Human polynucleoti
C 41	26	96.3	370	22	AAI08628	Human breast canc
C 42	26	96.3	373	16	AAI26624	Human gene signatu
C 43	26	96.3	373	22	AAAS33812	Human cDNA encodin
C 44	26	96.3	376	23	ABV13626	Human prostate exp
C 45	26	96.3	379	25	ABX33663	Bovine EST associa

ALIGNMENTS

RESULT 1

AAL31879/c
ID AAL31879 standard; DNA; 51 BP.
XX
AC AAL31879;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #5087.

Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
amyloid protein; angiotensin; apoptosis related protein; cadherin;
cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
interleukin; G-protein coupled receptor; thioesterase; inflammation;
multifactorial disease; autoimmune disease; infection;

KW nervous system disease; ss.
 XX Homo sapiens.
 OS WO200147944-A2.
 PN XX
 PD XX
 XX 05-JUL-2001.
 XX 28-DEC-2000; 2000WO-US35498.
 XX 28-DEC-1999; 99US-0173419.
 XX 27-DEC-2000; 2000US-0173419.
 XX (CURA-) CURAGEN CORP.
 PA Shimkets RA, Leach M;
 XX
 PI
 XX
 XX WPI; 2001-465210/50.
 DR
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 XX Claim 1; Page 2850; 4143pp; English.
 PS
 XX The present invention relates to oligonucleotides encoding polymorphic
 CC variants of proteins related to amylases, amyloid proteins, angiotensin,
 CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
 CC histones, kinases, colony stimulating factors, complement related
 CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
 CC G-protein coupled receptors and thioesterases. The present sequence is
 CC one such oligonucleotide. The oligonucleotides and the peptides encoded
 CC by them may be used in the prevention, diagnosis and treatment of
 CC diseases associated with inappropriate expression of the proteins listed
 CC above. Disorders that may be prevented, diagnosed and/or treated include
 CC multifactorial diseases with a genetic component, such as autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
 CC systemic lupus erythematosus and Grave's disease), inflammation, cancer
 CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
 CC leukaemia), diseases of the nervous system and an infection of pathogenic
 CC organisms.
 XX
 XX SQ Sequence 51 BP; 12 A; 5 C; 5 G; 29 T; 0 other;
 Alignment Scores:
 Pred. No.: 77.6 Length: 51
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 22 Gaps: 0
 US-09-492-764B-20 (1-6) x AAL31879 (1-51)
 QY 1 Pro***LeuLy5Thriys 6
 Db 48 CCATCTTTTAAAAACCAAA 31
 RESULT 2
 ABA71554/c
 ID ABA71554 standard; DNA; 162 BP.
 XX
 XX ABA71554;
 XX
 XX 01-FEB-2002 (first entry)
 XX Human foetal liver single exon nucleic acid probe #19859.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX Homo sapiens.
 OS
 PN WO200157277-A2.

XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00669.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-483447/52.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 PT
 XX Claim 4; SEQ ID NO 19859; 639pp + sequence listing; English.
 PS
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SQ Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other;
 Alignment Scores:
 Pred. No.: 283 Length: 162
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 22 Gaps: 0
 US-09-492-764B-20 (1-6) x ABA71554 (1-162)
 QY 1 Pro***LeuLy5Thriys 6
 Db 63 CCCTCTCTGAAARCCCAAG 46
 RESULT 3
 AAK19884/c
 ID AAK19884 standard; DNA; 162 BP.
 XX
 XX AAK19884;
 XX
 XX 05-NOV-2001 (first entry)
 XX Human brain expressed single exon probe SEQ ID NO: 19875.
 XX Human; brain expressed exon; gene expression analysis; probe;
 XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 XX epilepsy; cancer; ss.
 XX Homo sapiens.
 OS
 XX WO200157275-A2.
 PN
 PD 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00667.
 XX
 XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 19875; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system,
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other;

Alignment Scores:
 Pred. No.: 283 Length: 162
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 22 Gaps: 0

US-09-492-764B-20 (1-6) x AAK19884 (1-162)

QY 1 Pro**LeuLySThrLys 6
 Db 63 CCCTCTCTGAAAACCAAG 46
 RESULT 4
 AAK45914/c
 ID AAK45914 standard; DNA; 162 BP.
 XX AAK45914;
 XX 06-NOV-2001 (first entry)
 XX Human bone marrow expressed single exon probe SEQ ID NO: 20471.
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX Homo sapiens.
 XX WO200157276-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00668.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 19875; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system,
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other;

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 20471; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other;

Alignment Scores:
 Pred. No.: 283 Length: 162
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 22 Gaps: 0

US-09-492-764B-20 (1-6) x AAK45914 (1-162)

QY 1 Pro**LeuLySThrLys 6
 Db 63 CCCTCTCTGAAAACCAAG 46
 RESULT 5
 AAI51832/c
 ID AAI51832 standard; DNA; 162 BP.
 XX AAI51832;
 XX 17-OCT-2001 (first entry)
 XX Probe #20518 used to measure gene expression in human placenta sample.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX Homo sapiens.
 XX WO200157272-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00663.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID No 20518; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 XX

Note: The sequence information for this patent does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.

Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other; Alignment Scores: Pred. No.: 283 Length: 162 Score: 26.00 Matches: 5 Percent Similarity: 83.33% Conservative: 0 Best Local Similarity: 83.33% Mismatches: 1 Query Match: 96.30% Indels: 0 DB: 23 Gaps: 0

Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other; Alignment Scores: Pred. No.: 283 Length: 162 Score: 26.00 Matches: 5 Percent Similarity: 83.33% Conservative: 0 Best Local Similarity: 83.33% Mismatches: 1 Query Match: 96.30% Indels: 0 DB: 22 Gaps: 0

US-09-492-764B-20 (1-6) x ABS45612 (1-162)

US-09-492-764B-20 (1-6) x AAI51832 (1-162)

Qy 1 Pro**LeuLysThrLys 6
Db 63 CCCTCTCTGAAACCAAG 46

Qy 1 Pro**LeuLysThrLys 6
Db 63 CCCTCTCTGAAACCAAG 46

RESULT 7

RESULT 6

ABS20199/c ID ABS20199 standard; DNA; 162 BP. XX ABS20199; AC ABS20199; DT 19-AUG-2002 (first entry) XX Human genome-derived single exon probe ORF from lung SEQ ID No 20190. XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD; chronic obstructive pulmonary disease; interstitial lung disease; familial idiopathic pulmonary fibrosis; neurofibromatosis; tuberous sclerosis; Gaucher's disease; Niemann-Pick disease; Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis; pulmonary histiocytosis; lymphangiomyomatosis; Karagener syndrome; pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia; Primary ciliary dyskinesia; pulmonary hypertension; hyaline membrane disease; open reading frame; ORF.

WO200157273-A2. XX Homo sapiens. XX OS Homo sapiens. XX PN WO200186003-A2. XX PD 15-NOV-2001. XX PF 30-JAN-2001; 2001WO-US000665. XX PR 04-FEB-2000; 2000US-180312P. PR 26-MAY-2000; 2000US-207456P. PR 30-JUN-2000; 2000US-0608408. PR 03-AUG-2000; 2000US-0632366. PR 21-SEP-2000; 2000US-0234687. PR 27-SEP-2000; 2000US-0236359. PR 04-OCT-2000; 2000GB-0024263. XX PA (MOLE-) MOLECULAR DYNAMICS INC. XX PI Penn SG, Hanzel DK, Chen W, Rank DR; XX DR WPI; 2001-488898/53. XX PT Human genome-derived single exon nucleic acid probes useful for XX PS analysing gene expression in human adult liver - XX PS Claim 4; SEQ ID No 20602; 658pp; English.

Human liver single exon probe, SEQ ID No 20602. XX Human; single exon nucleic acid probe; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia; coronary heart disease; ss. XX OS Homo sapiens. XX PN WO200157273-A2. XX PD 09-AUG-2001. XX PF 30-JAN-2001; 2001WO-US000664. XX PR 04-FEB-2000; 2000US-0180312. PR 26-MAY-2000; 2000US-0207456. PR 30-JUN-2000; 2000US-0608408. PR 03-AUG-2000; 2000US-0632366. PR 21-SEP-2000; 2000US-0234687. PR 27-SEP-2000; 2000US-0236359. PR 04-OCT-2000; 2000GB-0024263. XX PA (MOLE-) MOLECULAR DYNAMICS INC. XX PI Penn SG, Hanzel DK, Chen W, Rank DR; XX DR WPI; 2001-488898/53. XX PT Human genome-derived single exon nucleic acid probes useful for XX PS analysing gene expression in human adult liver - XX PS Claim 4; SEQ ID No 20602; 658pp; English.

The invention relates to a single exon nucleic acid probe (SENP) (I) for measuring human gene expression in a sample derived from human adult liver, comprising one of 13109 defined nucleotide sequences given in the specification (or complements/ fragments). The probe hybridises at high stringency to a nucleic acid molecule expressed in the human adult liver. (II) may be used for predicting, measuring and displaying gene expression in samples derived from human adult liver. The genes identified may be involved in genetic liver diseases, such as cirrhosis, hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is associated with coronary heart disease. ABS25011-ABS51005 represent human liver single exon nucleic acid probes of the invention.

Spatially-addressable set of single exon nucleic acid probes, used to measure gene expression in human lung samples - Claim 4; SEQ ID No 20190; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA to a single exon probe
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic,
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 162 BP; 39 A; 39 C; 46 G; 38 T; 0 other;

Alignment Scores: Length: 162
 Pred. No.: 283 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 24 Gaps: 0

US-09-492-764B-20 (1-6) x ABS20199 (1-162)

Qy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 63 CCTCTCTGAAACCAAG 46

RESULT 8
 AAQ76958/c
 ID AAQ76958 standard; DNA; 183 BP.

XX AC AAQ76958;

XX DT 25-MAR-2003 (updated)

XX DT 23-SEP-1994 (first entry)

XX DE Human genome fragment.

XX KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
 KW detection; homology; human; adrenal tissue; ds.

XX OS Homo sapiens.

XX PN WO9401548-A2.

XX PD 20-JAN-1994.

XX PF 13-JUL-1993; 93WO-GB01467.

PR 13-JUL-1992; 92GB-0014857.
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX PA
 XX PI Cross J, Hadfield KM, Howells D, Kelly M, Shaw D;
 PI Sibson DR, Starkey M;
 DR WPI; 1994-035056/04.
 XX New nucleic acid fragment encoding gene products - can be used
 PT for genetic analysis and mapping

XX Claim 1; Page 290; 616pp; English.

XX Human nucleic acid fragments, isolated from brain adrenal tissue,
 CC the placenta or bone marrow comprise any of: (A) a sequence
 CC selected from (AAQ76401-077613), (B) an allelic variation of a
 CC sequence as described in (A), or (C) a sequence complementary
 CC to (A) or (B).

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 183 BP; 62 A; 27 C; 38 G; 56 T; 0 other;

Alignment Scores: Length: 183
 Pred. No.: 324 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 15 Gaps: 0

US-09-492-764B-20 (1-6) x AAQ76958 (1-183)

Qy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 134 CCTTCACTGAAACAAA 117

RESULT 9
 AAC18922/c
 ID AAC18922 standard; cDNA; 208 BP.

XX AC AAC18922;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 22997.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX PN BP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX DE (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 22997; 7ipp + CD-ROM; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

SQ Sequence 208 BP; 50 A; 45 C; 30 G; 83 T; 0 other;

Alignment Scores: Length: 208
 Pred. No.: 374 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 DB: 21 Gaps: 0

US-09-492-764B-20 (1-6) x AAC18922 (1-208)

QY 1 Pro**LeuLysThrLys 6
 ||| |||||
 Db 35 CCTAGTTTAAAGACAAG 18

RESULT 10
 AAX88587/c
 ID AAX88587 standard; DNA; 213 BP.
 XX
 AC AAX88587;
 XX
 DT 10-SEP-1999 (first entry)
 XX
 DE Human chromosome 18q YAC clone primer.
 XX
 KW Human chromosome 18q; mood disorder; polymorphic marker; detection;
 KW identification; trinucleotide repeat expansion; schizophrenia;
 KW anxiety disorder; adjustment disorder; personality disorder;
 KW nucleotide triplet repeat; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX W09932643-A2.
 XX
 XX 01-JUL-1999.
 XX
 XX 17-DEC-1998; 98WO-EP08543.
 XX
 XX 18-DEC-1997; 97GB-0026804.
 XX
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
 XX
 XX Del-Favero J, Raeymaekers P, Van Broeckhoven C;
 XX
 XX WPI; 1999-418934/35.
 XX
 XX Detecting nucleotide triplet repeats in human chromosome 18q
 XX
 XX Disclosure; Page 41; 87pp; English.
 XX

CC The present invention describes detecting nucleotide triplet repeats in
 CC a region of human chromosome 18q disposed between polymorphic markers
 CC D18S68 and D18S79 to identify a human gene associated with a mood
 CC disorder or related disorder. AAX88542 to AAX88705 represents human
 CC chromosome 18q YAC clones and primers corresponding to them, used in the
 CC exemplification of the present invention. YAC clones comprising a
 CC portion of the region of human chromosome 18q between markers D18S68 and

CC D18S79 are used to identify at least one human gene associated with a
 CC mood disorder or related disorder. The mood disorder or related
 CC disorder, is chosen from the Diagnostic and Statistical Manual of Mental
 CC Disorders, version 4 (DSM-IV) taxonomy. This includes mood disorders
 CC (296.XX, 300.4, 311, 301.13, 295.70), schizophrenia and related
 CC disorders (295, 297.1, 298.9, 297.3, 298.9), anxiety disorders (300.XX,
 CC 309.81, 308.13), adjustment disorders (309.XX) and personality disorders
 CC (codes 301.XX). Probes derived from genes associated with the mood
 CC disorder or related disorder can be used to detect pathological
 CC mutations or genetic variations in patients. The methods, probes and
 CC antibodies can be used to determine the susceptibility of an individual
 CC to a mood disorder or related disorder. The nucleic acids and proteins
 CC of the human gene can be used to treat mood disorders and related
 CC disorders.

SQ Sequence 213 BP; 64 A; 39 C; 38 G; 72 T; 0 other;

Alignment Scores: Length: 213
 Pred. No.: 384 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 DB: 20 Gaps: 0

US-09-492-764B-20 (1-6) x AAX88587 (1-213)

QY 1 Pro**LeuLysThrLys 6
 ||| |||||
 Db 113 CCATCTCTAAAACAAAA 96

RESULT 11
 ABZ79256/c
 ID ABZ79256 standard; DNA; 222 BP.
 XX
 AC ABZ79256;
 XX
 XX 24-APR-2003 (first entry)
 XX
 DE Tumour suppression-related sequence, SEQ ID 963.
 XX
 KW Cytostatic; virucide; apoptotic; gene therapy; tumour suppression;
 KW tumour reversion; apoptosis; virus resistance; viral infection; tumour;
 KW cell degenerative disease; ds.
 XX
 OS Unidentified.
 OS
 XX W0200264731-A2.
 XX
 XX 22-AUG-2002.
 XX
 XX 13-FEB-2002; 2002WO-FR00543.
 XX
 XX 13-FEB-2001; 2001FR-0001925.
 XX
 XX (MOLE-) MOLECULAR ENGINES LAB.
 XX
 XX Teerman A, Amson R, Tuijnder M, Susini L;
 XX
 XX WPI; 2003-058286/05.
 XX
 XX New nucleic acid encoding a translationally controlled tumor protein,
 XX useful for treating, preventing and diagnosing viral, tumor or
 XX degenerative diseases
 XX
 XX Disclosure; Page -; 45pp; French.
 XX

CC The present invention relates to novel nucleic acid sequences
 CC (ABZ78294-ABZ79313), which are involved in the molecular pathways of
 CC tumour suppression, tumour reversion, apoptosis and/or virus resistance.
 CC The sequences are also useful for treatment or prevention of viral,
 CC tumour and cell degenerative diseases, and also for diagnosis and
 CC prognosis of these diseases.

CC Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office.

XX SQ Sequence 222 BP; 61 A; 36 C; 42 G; 83 T; 0 other;

Alignment Scores: Pred. No.: 403 Length: 222 Score: 26.00 Matches: 5 Percent Similarity: 83.33% Conservative: 0 Best Local Similarity: 83.33% Mismatches: 1 Query Match: 96.30% Indels: 0 DB: 25 Gaps: 0

US-09-492-764B-20 (1-6) x ABZ79256 (1-222)

Qy 1 Pro***LeuLysThrLys 6
||| ||||||||
Db 214 CCATCTTTAAAAAACTAAA 197

RESULT 12
ABZ09803/C
ID ABZ09803 standard; DNA; 222 BP.
XX AC
XX AC ABZ09803;
XX DT 16-JAN-2003 (first entry)
XX DE Human oligonucleotide SEQ ID 963.

XX KW Human; tumour suppressor; virucide; cytostatic; nootropic; neuroprotective; neuroleptic; gene therapy; tumour suppression; tumour reversion; apoptosis; viral resistance; viral infection; cell degeneration; Alzheimer's disease; schizophrenia; cancer; ds.
XX OS Homo sapiens.
XX PI
XX PN FR2822475-A1.
XX PD 27-SEP-2002.
XX PF 20-MAR-2002; 2002FR-0003459.
XX PR 13-FEB-2001; 2001FR-0001925.
XX PA (MOLE-) MOLECULAR ENGINES LAB SA.

XX PI Teierman A, Amson R, Tuijnder M, Susini L;
XX DR WPI; 2003-032204/03.
XX PT New human nucleic acid, useful for diagnosis, prognosis and treatment, e.g. of tumors, also related vectors, transformed cell, polypeptides and antibodies -
XX PS Disclosure; Page 180; 189pp; French.
XX CC The present invention relates to human oligonucleotides (ABZ08841-ABZ09860). The expression of the oligonucleotides is implicated in tumour suppression or reversion, apoptosis and/or viral resistance.
XX CC The oligonucleotides are useful for preventing and/or treating viral infection, tumour development and cell degeneration (e.g. Alzheimer's disease and schizophrenia), especially cancer.

XX SQ Sequence 222 BP; 61 A; 36 C; 42 G; 83 T; 0 other;

Alignment Scores: Pred. No.: 403 Length: 222 Score: 26.00 Matches: 5 Percent Similarity: 83.33% Conservative: 0 Best Local Similarity: 83.33% Mismatches: 1 Query Match: 96.30% Indels: 0 DB: 25 Gaps: 0

US-09-492-764B-20 (1-6) x ABZ09803 (1-222)

Qy 1 Pro***LeuLysThrLys 6
||| ||||||||
Db 214 CCATCTTTAAAAAACTAAA 197

RESULT 13
ABZ78118/C
ID ABZ78118 standard; cDNA; 227 BP.
XX AC
XX AC ABZ78118;
XX DT 19-MAY-2003 (first entry)
XX DE Human suppression subtractive hybridisation cDNA fragment #17.

XX KW Human; cytostatic; vaccine; cancer; immune response; suppression subtractive hybridisation; SSH; ss.
XX OS Homo sapiens.
XX PN WO200283921-A2.
XX PD 24-OCT-2002.
XX PF 10-APR-2002; 2002WO-US11654.
XX PR 10-APR-2001; 2001US-282739P.
XX PR 10-APR-2001; 2001US-283112P.
XX PR 25-APR-2001; 2001US-286630P.
XX PA (AGEN-) AGENSYS INC.

XX PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS,
XX PI Morrison K, Morrison RK, Raitano AB;
XX DR WPI; 2003-075555/07.
XX PT New composition comprising a substance that modulates the structure of proteins and polynucleotides, useful for therapeutic, prognostic and diagnostic reagents for eliciting cellular or humoral immune response in cancer patients -
XX PS Example 1; Fig 1Q; 1021pp; English.
XX CC The present invention relates to novel human cancer-related genes and proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and proteins are useful for eliciting a humoral or cellular immune response.
XX CC The genes are useful as probes and primers for the amplification and/or detection of genes, mRNAs or their fragments, as reagents for the diagnosis and/or prognosis of cancer, as coding sequences capable of directing the expression of the protein, as tools for modulating or inhibiting the expression of genes and/or translation of transcripts, and as therapeutic agents. The proteins and peptides are useful as therapeutic, prognostic and diagnostic reagents for cancer. The present sequence is a suppression subtractive hybridisation (SSH) cDNA fragment used in an example from the invention.

XX SQ Sequence 227 BP; 53 A; 56 C; 51 G; 67 T; 0 other;

Alignment Scores: Pred. No.: 413 Length: 227 Score: 26.00 Matches: 5 Percent Similarity: 83.33% Conservative: 0 Best Local Similarity: 83.33% Mismatches: 1 Query Match: 96.30% Indels: 0 DB: 25 Gaps: 0

US-09-492-764B-20 (1-6) x ABZ78118 (1-227)

Qy 1 Pro***LeuLysThrLys 6
||| ||||||||

DE Clone G31C6B cDNA fragment.
 XX Human; molecular marker; cervical cancer tissue; clone G31C6B; ss.
 XW Homo sapiens.
 XX WO200208419-A1.
 XX 31-JAN-2002.
 XX 26-JUL-2001; 2001WO-AU00910.
 XX 26-JUL-2000; 2000AU-0009017.
 XX (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
 XX (HUGH/) HUGHES E J L.
 XX Hui KM, Cheng Q;
 XX WPI; 2002-227050/28.
 XX Detecting molecular marker associated with normal or abnormal tissue
 XX e.g. from cervical carcinoma of a mammal, e.g. human by comparing
 XX differential expression of genetic sequences in cancer relative to
 XX normal tissue
 XX Claim 46; Page 43; 80pp; English.
 XX The present invention relates to a method for detecting molecular marker
 XX associated with normal or abnormal tissue from mammal. The method
 XX comprising generating cDNA from total RNA isolated from abnormal tissue
 XX and corresponding normal tissue from normal individual, and separating
 XX the cDNA so that the presence or absence of the cDNA can be detected from
 XX abnormal tissue relative to normal tissue. The presence or absence of
 XX the cDNA indicates of molecular marker associated with normal or abnormal
 XX tissue. The method is useful for detecting a molecular marker associated
 XX with normal or abnormal tissue from a mammal, preferably human, where
 XX abnormal tissue is cervical cancer tissue. Isolated genetic sequence is
 XX useful in distinguishing between normal tissue and cervical cancer tissue
 XX and the isolated primer is useful in RT-PCR of the method. The
 XX identification of differentially expressed genetic sequence which is up-
 XX or down-regulated in cancer cells relative to normal cells provides
 XX useful diagnostic aid. The present sequence is a human cDNA fragment used
 XX in the invention.
 XX SQ Sequence 243 BP; 78 A; 36 C; 48 G; 81 T; 0 other;

DE Clonage de l'ADN de la tumeur humaine.
 XX Clonage de l'ADN de la tumeur humaine; tissu cancéreux du col de l'utérus; séquence de l'ADN G31C6B; ss.
 XW Homo sapiens.
 XX WO200208419-A1.
 XX 31-JAN-2002.
 XX 26-JUL-2001; 2001WO-AU00910.
 XX 26-JUL-2000; 2000AU-0009017.
 XX (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
 XX (HUGH/) HUGHES E J L.
 XX Hui KM, Cheng Q;
 XX WPI; 2002-227050/28.
 XX Détection d'un marqueur moléculaire associé à un tissu normal ou anormal
 XX tel que, par exemple, un carcinome du col de l'utérus d'un mammifère, en
 XX comparant l'expression différentielle de séquences génétiques dans le tissu
 XX cancéreux par rapport au tissu normal.
 XX Revendication 46; Page 43; 80 pages; anglais.
 XX La présente invention se rapporte à une méthode pour détecter un marqueur
 XX moléculaire associé à un tissu normal ou anormal issu d'un mammifère. La
 XX méthode comprend la génération d'ADNc à partir d'ARN total isolé à partir
 XX d'un tissu anormal et d'un tissu normal issu d'un individu normal, et la
 XX séparation de l'ADNc de telle sorte que la présence ou l'absence de l'ADNc
 XX peut être détectée dans le tissu anormal par rapport au tissu normal.
 XX La présence ou l'absence de l'ADNc indique la présence ou l'absence d'un
 XX marqueur moléculaire associé à un tissu normal ou anormal. La méthode est
 XX utile pour détecter un marqueur moléculaire associé à un tissu normal ou
 XX anormal issu d'un mammifère, de préférence humain, dans lequel le tissu
 XX anormal est un tissu cancéreux du col de l'utérus. Une séquence génétique
 XX isolée est utile pour distinguer entre un tissu normal et un tissu cancéreux
 XX du col de l'utérus, et le primer isolé est utile pour la RT-PCR de la
 XX méthode. L'identification de séquences génétiques exprimées différemment
 XX qui sont surexprimées ou sous-exprimées dans les cellules cancéreuses
 XX par rapport aux cellules normales fournit une aide diagnostique utile. La
 XX présente séquence est un fragment d'ADN humain utilisé dans l'invention.
 XX SQ Séquence de 243 pb; 78 A; 36 C; 48 G; 81 T; 0 autres;

Alignment Scores:
 Pred. No.: 446 Length: 243
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 24 Gaps: 0
 US-09-492-764B-20 (1-6) x AAD30827 (1-243)
 QY 1 Pro***LeulystThrls 6
 DB 173 CCCTCATTGAAACAAAA 156

Alignment Scores:
 Pred. No.: 446 Length: 243
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 24 Gaps: 0
 US-09-492-764B-20 (1-6) x ABT06899 (1-243)
 QY 1 Pro***LeulystThrls 6
 DB 145 CCAGCCTTAAGACCAAG 162

Search completed: August 7, 2003, 17:58:35
 Job time : 179 secs

Search completed: August 7, 2003, 17:58:35
 Job time : 179 secs

RESULT 14
 ID ABT06899 standard; cDNA; 243 BP.
 AC ABT06899;
 XX 07-NOV-2002 (first entry)
 XX Human ovarian cancer associated coding sequence SEQ ID NO: 61.
 XX Human; ovarian cancer; cancer; gene; ss.
 XX Homo sapiens.
 XX US2002076715-A1.
 XX 20-JUN-2002.
 XX 06-JUN-2001; 2001US-0876889.
 XX 23-SEP-1998; 98US-0159320.
 XX 08-FEB-1999; 99US-0246429.
 XX 16-SEP-1999; 99US-0397787.
 XX (BENS/) BENSON D R.
 XX (LODE/) LODES M J.
 XX (MITC/) MITCHAM J L.
 XX (KING/) KING G E.
 XX Benson DR, Lodes MJ, Mitcham JL, King GE;
 XX WPI; 2002-598720/64.
 XX Composition for detecting and treating ovarian cancer, comprises a
 XX specific polypeptide, polynucleotide, T cell population, or antigen
 XX presenting cell
 XX Example 1; Fig 16; 188pp; English.
 XX The present invention relates to a method of detecting the presence of
 XX ovarian cancer in a patient, involving detecting ovarian cancer
 XX associated polynucleotides. The method is not only used to detect the
 XX presence of cancer, preferably ovarian cancer in a patient, but also is
 XX used to stimulate and/or expand T cells specific for an ovarian tumour
 XX protein. The sequences can be used in vaccines used to treat cancer. The
 XX present sequence is an ovarian cancer associated coding sequence.

Alignment Scores:
 Pred. No.: 446 Length: 243
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 24 Gaps: 0
 US-09-492-764B-20 (1-6) x AAD30827 (1-243)
 QY 1 Pro***LeulystThrls 6
 DB 145 CCAGCCTTAAGACCAAG 162

RESULT 15
 AAD30827/c
 ID AAD30827 standard; cDNA; 243 BP.
 XX AAD30827;
 XX 31-MAY-2002 (first entry)

RESULT 15
 AAD30827/c
 ID AAD30827 standard; cDNA; 243 BP.
 XX AAD30827;
 XX 31-MAY-2002 (first entry)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 9, 2003, 18:20:24 ; Search time 1546 Seconds
(without alignments)
94.325 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLTKT 6

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392_seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-C=/cgn2_1/USPTO_spool/US09492764/runat_04082003_142327_11138/app_query.fasta_1.199
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HGAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09492764@cgn_1_1_2135@runat_04082003_142327_11138 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssi:*

29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	26	96.3	88	28	BH227102	1006137C0
C 2	26	96.3	104	14	CD022154	NXPV_016
C 3	26	96.3	111	28	AZ037377	RPCI-23-3
C 4	26	96.3	113	28	BH813312	SALK_0639
C 5	26	96.3	116	10	BF909108	MR0-UT004
C 6	26	96.3	118	29	AL937043	Arabidops
C 7	26	96.3	121	9	AI909373	QV-BT204-
C 8	26	96.3	123	10	BF173306	MEY1175 M
C 9	26	96.3	127	9	AW454051	zeh10867
C 10	26	96.3	128	9	AI906850	RC-BT126-
C 11	26	96.3	130	12	BM196630	C0333607-
C 12	26	96.3	134	9	AW293049	Ur-H-Biz-
C 13	26	96.3	135	9	AW793735	MR1-UM000
C 14	26	96.3	141	9	AW793767	MR1-UM000
C 15	26	96.3	142	14	R28953	F0-2D 22 we
C 16	26	96.3	144	10	BE089177	CM2-BT069
C 17	26	96.3	146	28	BH270649	CH230-176
C 18	26	96.3	150	9	AW214348	uolnhil.x
C 19	26	96.3	152	9	AL718179	AL718179
C 20	26	96.3	153	12	BM205279	CO281D08-
C 21	26	96.3	154	29	CC326808	RR1159 Ba
C 22	26	96.3	156	9	AA281880	zt10h02.x
C 23	26	96.3	156	14	CD023493	NXPV_133
C 24	26	96.3	157	10	BG290356	602385988
C 25	26	96.3	158	9	AW580600	CM3-HT048
C 26	26	96.3	158	9	AW580601	CM3-HT048
C 27	26	96.3	159	9	AW384034	RC1-HT037
C 28	26	96.3	160	28	AZ786313	2M0031D15
C 29	26	96.3	162	14	R83034	YP87G02.81
C 30	26	96.3	163	28	AZ596889	IM041UM21
C 31	26	96.3	165	14	HSM091139	Hom0_sap1
C 32	26	96.3	165	14	CB116259	K-EST0160
C 33	26	96.3	166	9	AW800841	MR3-UM006
C 34	26	96.3	166	13	BU952874	i188a09.x
C 35	26	96.3	167	9	AV098914	AV098914
C 36	26	96.3	168	13	BU651803	1112095F1
C 37	26	96.3	170	28	AZ255157	RPCI-23-1
C 38	26	96.3	171	28	AZ077341	RPCI-23-3
C 39	26	96.3	174	28	AZ833030	2M0113M17
C 40	26	96.3	175	28	AZ108660	RPCI-23-3
C 41	26	96.3	176	14	CA951198	ie47a12.x
C 42	26	96.3	177	10	BF462331	Ut-M-CG0P
C 43	26	96.3	177	28	AZ799738	2M0057B24
C 44	26	96.3	180	14	N50278	yy84h04.r1
C 45	26	96.3	182	9	AA234329	zr71f02.r

ALIGNMENTS

RESULT 1
BH227102/c
LOCUS BH227102
DEFINITION 1006137C02.2EL_y1 1006 - Rescuemu Grid G Zea mays genomic, genomic survey sequence.
ACCESSION BH227102.1 GI:16826757
VERSION GSS.
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays

BH227102 88 bp DNA linear GSS 08-NOV-2001
1006137C02.2EL_y1 1006 - Rescuemu Grid G Zea mays genomic, genomic survey sequence.
BH227102.1 GI:16826757
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

```
REFERENCE
AUTHORS      1 (bases 1 to 88)
TITLE        Walbot.V.
JOURNAL      Maize genomic sequences found using engineered RescueMu transposon
COMMENT      Unpublished
              Contact: Walbot V
              Department of Biological Sciences
              Stanford University
              855 California Ave. Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Possible ligation site of ends cut by 2 different endonucleases.
              Reverse complemented post-ligation sequence from source sequence.
              Plate: 1006137 row: 21
              Class: transposon-tagged.
```

```
FEATURES
source
1. .88
   /organism="Zea mays"
   /mol_type="genomic DNA"
   /cultivar="mixed background W23/A188/B73"
   /db_xref="taxon:4577"
   /tissue_type="leaf"
   /dev_stage="adult"
   /lab_host="DH10B"
   /clone_lib="1006 - RescueMu Grid G"
   /note="Organ: leaf; Vector: RescueMu (engineered from pBluescript backbone); Site 1: BamHI; Site 2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu', 'Grid G' was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."
```

```
BASE COUNT 17 a 29 c 27 g 14 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3.66e+03 Length: 88
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 28 Gaps: 0
```

```
US-09-492-764B-20 (1-6) x BH227102 (1-88)
Oy 1 Pro***LeuLysThrLys 6
Db 35 CCTTCCTTTGAAGACTAAA 18
```

```
RESULT 2
CD022154 104 bp mRNA linear EST: 07-MAY-2003
LOCUS NXPV_016_B02_F_NXPV (Nsf Xylem Planings wood Vertical) Pinus taeda
DEFINITION cDNA clone NXPV_016_B02_5', mRNA sequence.
VERSION CD022154
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
REFERENCE 1 (bases 1 to 104) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS Sederoff, R.
TITLE Molecular Basis of Wood Formation in the Pine Megagenome
COMMENT Unpublished
          Contact: Sederoff, Ron
          Forest Biotechnology
          North Carolina State University
```

```
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,
NC 27695, USA
Tel: 919 515 7800
Fax: 919 515 7801
Email: ron_sederoff@ncsu.edu
Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further
information.
Seq primer: T3.
```

```
FEATURES
source
Location/Qualifiers
1. .104
   /organism="Pinus taeda"
   /mol_type="mRNA"
   /strain="Coastal plain loblolly pine from North Carolina"
   /db_xref="taxon:3352"
   /clone="NXPV_016_B02"
   /tissue_type="Xylem"
   /cell_type="planings (secondary)"
   /dev_stage="transitional"
   /lab_host="X11-Blue"
   /clone_lib="NXPV (Nsf Xylem Planings wood Vertical)"
   /note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) secondary wood, taken
from a ten year old tree in the transitional phase. The
tree is a kind gift of the Westvaco Corporation. Secondary
xylem was harvested from the tree by peeling back the bark
and primary xylem and then removing the underlying tissue
with a block plane. NOTE: The sequences contain a 'cDNA
adapter' between the EcoRI site and the start of the EST.
The adapter sequence is 'AAATTCGGCAGG'."
BASE COUNT 24 a 20 c 24 g 32 t 4 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 4.36e+03 Length: 104
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 14 Gaps: 0
```

```
US-09-492-764B-20 (1-6) x CD022154 (1-104)
Oy 1 Pro***LeuLysThrLys 6
Db 15 CCTGCATTAAAGACGAAA 32
```

```
RESULT 3
AZ037377 111 bp DNA linear GSS 01-MAR-2000
LOCUS RPCI-23-364L5.TJ RPCI-23 Mus musculus genomic clone RPCI-23-364L5,
DEFINITION genomic survey sequence.
VERSION AZ037377 GI:7125780
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 111)
          Zhao, S., Niernan, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
          B., LeVins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
          and Fraser, C.M.
          Mouse BAC End Sequences from Library RPCI-23
          Unpublished
          Other GSSs: RPCI-23-364L5.TV
          Contact: Shaying Zhao
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: szhao@tigr.org
          Clones are derived from the mouse BAC library RPCI-23. For BAC
```

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
 Plate: 364 row: L column: 5
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Location/Qualifiers

1..111
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-364L5"
 /sex="Female"
 /lab_host="DH10B"
 /clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1: ECORI; Site_2: ECORI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRII. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 36 a 31 c 16 g 28 t

BASE COUNT

36 a 31 c 16 g 28 t

Alignment Scores:

Pred. No.: 4.67e+03 Length: 111
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 28 Gaps: 0

US-09-492-764B-20 (1-6) x AZ037377 (1-111)

Qy 1 Pro**LeuLysThrLys 6
 ||| |||||
 66 CCACGCTTTAAAACAAA 83

RESULT 4

BH813312/c BH813312 113 bp DNA linear GSS 02-MAY-2002
 LOCUS SALK_063966 Arabidopsis thaliana TDNA insertion lines Arabidopsis
 DEFINITION thaliana genomic clone SALK_063966, genomic survey sequence.

ACCESSION

BH813312 GI:20391767

KEYWORDS

Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsi

REFERENCE

1 (bases 1 to 113)
 Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
 ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
 , Zimmerman,J. and Ecker,J.R.

TITLE

A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome

JOURNAL

Unpublished
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)

COMMENT

The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: ecker@salk.edu
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within an annotated exon of At5g94420.
 Class: TDNA tagged.

FEATURES

Location/Qualifiers

source

1..113
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_063966"
 /clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"
 40 a 11 c 23 g 37 t

BASE COUNT

40 a 11 c 23 g 37 t

ORIGIN

Alignment Scores:
 Pred. No.: 4.76e+03 Length: 113
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 28 Gaps: 0

US-09-492-764B-20 (1-6) x BH813312 (1-113)

Qy 1 Pro**LeuLysThrLys 6
 ||| |||||
 52 CCITCATTTAAAACAAA 35

RESULT 5

BF909108 116 bp mRNA linear EST 18-JAN-2001
 LOCUS WR0-UT0049-151000-104-a03 UT0049 Homo sapiens cDNA, mRNA sequence.
 DEFINITION

ACCESSION

BF909108 GI:12300566

VERSION

EST.

KEYWORDS

Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 116)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663
 10737800

PUBMED

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research

COMMENT

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR06&t2=MR0-UT0049-151000-104-a03&t3=2000-10-15&t4=1>)

Seq primer:

puc 18 forward
 High quality sequence start: 17
 High quality sequence stop: 116.

FEATURES

Location/Qualifiers
 1..116
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 47 a 18 c 20 g 33 t

ORIGIN

Alignment Scores: Length: 118
 Pred. NO.: 4.99e+03
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 DB: 29

US-09-492-764B-20 (1-6) x AL937043 (1-118)

Oy 1 Pro**LeulysThrLys 6
 Db 10 CCTTCTTGAAGACTAAA 27

RESULT 7

AL909373/c 121 bp mRNA linear EST 30-MAR-2000
 LOCUS QV-BT204-060499-027 BT204 Homo sapiens cDNA, mRNA sequence.

DEFINITION AI909373
 ACCESSION AI909373
 VERSION AI909373.1 GI:6500053
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 121)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. K., Reis, L. F., de Souza, S. J. and Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/seq/gethtml.pl?tl=QV&t2=QV-BT204-027.html>)

Seq primer: puc 18 forward.
 Location/Qualifiers

1. .118
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="female"
 /dev_stage="Adult"
 /clone_lib="BT204"
 /notes="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 30 a 30 c 27 g 29 t
 ORIGIN

Alignment Scores: Length: 116
 Pred. No.: 4.9e+03
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 DB: 10

US-09-492-764B-20 (1-6) x BF909108 (1-116)

Oy 1 Pro**LeulysThrLys 6
 Db 18 CCTGCTCTTAAACAAAA 35

RESULT 6

AL937043 118 bp DNA linear GSS 23-OCT-2002
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-063D10-016170,

DEFINITION Arabidopsis thaliana (chale cress)
 ACCESSION AL937043
 VERSION AL937043.1 GI:24368668

KEYWORDS GSS.
 SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana (chale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 AUTHORS Strizhov, N., Li, Y., Rosso, M., Viehoever, P., Dekker, K., Saedler, H. and Weissshaar, B.

A pipeline for automated high-throughput generation of ESTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

TITLE Unpublished

JOURNAL REFERENCE
 AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weissshaar, B.
 TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

JOURNAL REFERENCE
 AUTHORS Rosso, M., Li, Y., Strizhov, N. and Weissshaar, B.

Submitted (21-OCT-2002) Weissshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln 50829, Germany

This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g41060. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:

http://www.mpiz-koeln.mpg.de/GABI-Kat/

Location/Qualifiers

1. .118
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone_lib="GK-063D10-016170"
 /notes="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from

BASE COUNT 42 a 20 c 19 g 40 t
 ORIGIN

size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing."

Alignment Scores: 5.12e+03 Length: 121
 Pred. No.: 26.00 Matches: 5
 Score: 83.33% Conservatvie: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 96.30% Indels: 0
 Query Match: 9 Indels: 0
 DB: Gaps: 0

US-09-492-764B-20 (1-6) x AI909373 (1-121)
 Qy 1 Pro**LeuLystrLys 6
 Db 111 CCTGCCCTGAAACCAAA 94

RESULT 8
 BF173306/c BF173306 123 bp mRNA linear EST 23-MAR-2001
 LOCUS MYEL175 Myeloma (MYE) cDNA library Homo sapiens CDNA, mRNA
 DEFINITION sequence.

ACCESSION BF173306
 VERSION BF173306.1 GI:13439520
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 123)
 AUTHORS Claudio, J.O., Masih-Khan, E., Tang, H., Goncalves, J., Voralia, M., Li, Z.H., Nadeem, V., Cukerman, E., Francisco-Pabalan, O., Liew, C.C., Woodgett, J.R., and Stewart, A.K.
 TITLE A molecular compendium of genes expressed in multiple myeloma
 JOURNAL Blood 100 (6), 2175-2186 (2002)
 MEDLINE 22188429
 PUBMED 12200383

COMMENT Contact: A. Keith Stewart, M.D.
 Oncology Research
 University Health Network
 610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
 Tel: (416) 946-4639
 Fax: (416) 946-6546
 Email: k.stewart@utoronto.ca
 PCR Primers
 FORWARD: 5'-GCCAAGCTCGAAATTAACCCCTCACTAAAGG-3'
 BACKWARD: 5'-CCAGTGAATTGTAATACGACTACTATAGGGCG-3'
 Seq primer: 5'-GAAATTAACCCCTCACTAAAGG-3'

FEATURES
 source
 1..123
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="male"
 /tissue="Blood"
 /cell_type="myeloma"
 /dev_stage="multiple myeloma"
 /clone_lib="Myeloma (MYE) cDNA library"
 /note="Vector: Lambda Zap Express; Site 1: EcoRI; Site 2: XhoI; Myeloma cells from multiple myeloma patients' bone marrow were purified by magnetic cell sorting. mRNA were purified and an oligo (dT)18 primer containing XhoI restriction site was used to prime first strand synthesis using M-MLV reverse transcriptase. To protect the cDNAs from XhoI digestion in subsequent cloning step, the nucleotide analogue 5-methyl-dCTP was added to the nucleotide mixture and [a-32P]dATP was added to monitor the quantity and quality of first strand synthesis. After second-strand synthesis and blunting of cDNA termini, EcoRI adapters were ligated, followed by kinase treatment and digestion with XhoI. The cDNAs were then

BASE COUNT 23 a 33 c 35 g 32 t
 ORIGIN

size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing."

Alignment Scores: 5.21e+03 Length: 123
 Pred. No.: 26.00 Matches: 5
 Score: 83.33% Conservatvie: 0
 Percent Similarity: 83.33% Mismatches: 1
 Best Local Similarity: 96.30% Indels: 0
 Query Match: 9 Indels: 0
 DB: Gaps: 0

US-09-492-764B-20 (1-6) x BF173306 (1-123)
 Qy 1 Pro**LeuLystrLys 6
 Db 101 CCGTCCTTAAACCAACCAAG 84

RESULT 9
 AW454051/c AW454051 127 bp mRNA linear EST 17-FEB-2000
 LOCUS Zeh10867 Zebrafish Embryonic Heart CDNA Library Danio rerio CDNA
 DEFINITION 5' mRNA sequence.

ACCESSION AW454051
 VERSION AW454051.1 GI:6994837
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 127)
 AUTHORS Ton, C., Mably, J.D., Dempsey, A.A., Hwang, D.M., Fishman, M.C. and Liew, C.C.
 TITLE Identification and Characterization of Expressed Sequence Tags from an Embryonic Zebrafish Heart CDNA Library
 JOURNAL Unpublished
 COMMENT Contact: Liew CC
 Brigham and Women's Hospital
 Harvard Medical School
 75 Francis St., Boston, MA 02115, USA
 Tel: 6177328915
 Fax: 6179750995
 Email: cliew@rics.bwh.harvard.edu
 PCR Primers
 FORWARD: 5' GCCAAGCTCGAAATTAACCCCTCACTAAAGG 3'
 BACKWARD: 5' CCAGTGAATTGTAATACGACTACTATAGGGCG 3'
 Seq primer: 5' GAAATTAACCCCTCACTAAAGG 3'

FEATURES
 Location/Qualifiers
 1..127
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /dev_stage="embryonic day 3 post-fertilization"
 /lab_host="E.coli XLI-Blue mrf."
 /clone_lib="Zebrafish Embryonic Heart CDNA Library"
 /note="Organ: heart; Vector: Lambda ZAP Express; Site 1: EcoRI; Site 2: XhoI; mRNA was purified from embryonic zebrafish hearts (3 day post-fertilization) cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into pre-digested lambda ZAP Express vector."

BASE COUNT 34 a 21 c 42 g 30 t
 ORIGIN

Alignment Scores: 5.39e+03 Length: 127
 Pred. No.:

Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 9 Gaps: 0

US-09-492-764B-20 (1-6) x AW454051 (1-127)

QY 1 Pro**Leu1ystrHlys 6
 DB 96 CCAAGCTGAAGACAAA 79

RESULT 10
 AI906850 128 bp mRNA linear EST 30-MAR-2000
 LOCUS RC-BT126-040399-043 BT126 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION AI906850
 VERSION AI906850.1 GI:6497258
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/seq/gethtml.pl?tl=RC&t2=RC-BT126-043.html
 &t3=040399&t4=1)

FEATURES
 source
 1..128
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="female"
 /dev_stage="Adult"
 /clone_lib="BT126"
 /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 from ORFESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 34 a 38 c 27 g 29 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.43e+03 Length: 128
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 9 Gaps: 0

US-09-492-764B-20 (1-6) x AI906850 (1-128)

QY 1 Pro**Leu1ystrHlys 6
 DB 36 CCTGCACCTAAGACAAAG 53

RESULT 11
 BM196630 130 bp mRNA linear EST 08-JUN-2003
 LOCUS C0333G07-3 NIA Mouse Undifferentiated ES Cell cDNA Library (long)
 DEFINITION Mus musculus cDNA clone NIA:C0333G07 IMAGE:30008046 3', mRNA
 sequence.
 ACCESSION BM196630
 VERSION BM196630.2 GI:31523083
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scaurognathi; Muridae; Murinae; Mus.
 AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 TITLE Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
 MEDLINE 21429098
 PUBMED 11544199
 COMMENT On Dec 14, 2001 this sequence version replaced gi:17748519.
 Other ESTs: C0333G07-5N
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: C0333 row: G column: 07
 Seq primer: -21M13 Forward
 High quality sequence stop: 130
 POLYA=Yes.

FEATURES
 source
 1..130
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x 129/Sv-Cp"
 /db_xref="niaEST:C0333G07-3"
 /db_xref="taxon:10090"
 /clone="NIA:C0333G07 IMAGE:30008046"
 /tissue_type="Undifferentiated ES Cell"
 /cell_line="R1 ES cells"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Undifferentiated ES Cell cDNA
 Library (Long)"
 /note="vector: pSPORT1 (Invitrogen); Site_1: Sali; Site_2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a
 long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
 obtained from Dr. Kenneth R. Boheler (National Institute
 on Aging, USA). ES cells were cultured without feeder
 cells in the presence of LIF and BRU-conditioned media.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen]:
 5'-pGACTAGTCTAGATCGCGCCCTTTTTTTTTTTTTTTT-3' from
 14.2 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lone-linker L1-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with Sali and NotI enzymes
 and cloned into Sali/NotI site of pSPORT1 plasmid vector.

The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.4 kb. The library was constructed by Yulan Piao (NIA)."

BASE COUNT 50 a 26 c 25 g 29 t
ORIGIN

Alignment Scores: 5.52e+03 Length: 130
Pred. No.: 26.00 Matches: 5
Score: 83.33% Conservat: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.30% Indels: 0
Query Match: 12 Gaps: 0

US-09-492-764B-20 (1-6) x BML96630 (1-130)

QY 1 Pro***LeuLyThrlYs 6
||| ||||| ||||| |||||
DB 41 CCATCTTGAAAACATAAA 58

RESULT 12
AW293049/c 134 bp mRNA linear EST 16-JAN-2000
LOCUS UI-H-B12-8gz-a-04-0-UI.e1 NCI_CGAP_Sub4 Homo sapiens cDNA clone
DEFINITION IMAGE:2725903 3', mRNA sequence.

ACCESSION AW293049 GI:6699685
VERSION AW293049.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 134)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapps-r@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image.html The following repetitive elements were found in this cDNA sequence: 15-123, Seq primer: M13 Forward POLYA=Yes.

FEATURES
source
1..134 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2725903"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI_CGAP Sub4 library is a subtracted library derived from the NCI_CGAP Sub2 library which is a subtracted library derived from the NCI_CGAP Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_LL11, NCI_CGAP_Le12, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_G4, NCI_CGAP_G6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture

was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 : LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones) 1322376-1322911, 1456008-1456775, 1500552-1502855) NCI_CGAP_Kid5 pool 1 : LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones) 1323912-1325831, 1471368-1472903, 1492104-1493255) NCI_CGAP_Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE Clones) 1414920-1417991, 1520904-1522439) NCI_CGAP_G4 pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones) 1257096-1258631, 1469064-1470983, 1475592-1476743) NCI_CGAP_Pr22 pool 1 : LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones) 985608-986759, 1101192-1101959, 1217928-1220615) NCI_CGAP_Co10 pool 1 : LLAM 2644-2653, 2871-2872 (IMAGE Clones) 1057416-1061255, 1144584-1144531)
Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.)

BASE COUNT 20 a 39 c 27 g 48 t
ORIGIN
TAG LIB=NCI_CGAP_Co10
TAG_TISSUE=Colon
TAG_SEQ=AAACG"

Alignment Scores: 5.7e+03 Length: 134
Pred. No.: 26.00 Matches: 5
Score: 83.33% Conservat: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 96.30% Indels: 0
Query Match: 9 Gaps: 0

US-09-492-764B-20 (1-6) x AW293049 (1-134)

QY 1 Pro***LeuLyThrlYs 6
||| ||||| ||||| |||||
DB 31 CCTAGTCTCAAAACAAAA 14

RESULT 13
AW93735
LOCUS MRI-UM0009-290200-001-h10 UM0009 Homo sapiens cDNA, mRNA sequence.
DEFINITION AW93735
ACCESSION AW93735
VERSION AW93735.1 GI:7845605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 135)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A. J. G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-MRI-UM0009-290)

```

source
1. .141
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0009"
/notes="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 46 a 38 c 19 g 38 t
ORIGIN
Alignment Scores: Length: 141
Pred. No.: 6.02e+03 Matches: 5
Score: 26.00 Conservatave: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 9
US-09-492-764B-20 (1-6) x AW793767 (1-141)
QY 1 Pro**LeuLysThrLys 6
||| ||||| |||||
Db 86 CCATCTTAAACTAAA 103

```

```

source
1. R35
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="UM0009"
/notes="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 45 a 33 c 19 g 38 t
ORIGIN
Alignment Scores: Length: 135
Pred. No.: 5.75e+03 Matches: 5
Score: 26.00 Conservatave: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 9
US-09-492-764B-20 (1-6) x AW793735 (1-135)
QY 1 Pro**LeuLysThrLys 6
||| ||||| |||||
Db 80 CCATCTTAAACTAAA 97

```

```

RESULT 15
R28953/c
LOCUS
DEFINITION
R28953 142 bp mRNA linear EST 03-DEC-1999
F0-2D 22 week old human fetal liver cDNA library Homo sapiens cDNA
clone F0-2D 5', mRNA sequence.
R28953 GI:1546936
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 142)
Choi, S.S., Yun, J.W., Choi, E.K., Cho, Y.G., Sung, Y.C. and Shin, H.S.
Construction of a gene expression profile of a human fetal liver by
single-pass cDNA sequencing
Mamm. Genome 6 (9), 653-657 (1995)
96081342
JOURNAL
MEDLINE
PUBMED
8535075
COMMENT
Contact: Hee-Sup Shin
Developmental Genetics
Pohang Institute of Science & Technology
Sanli, Hyojadong Pohang, 790-784 Republic of Korea
Tel: 562-279-2291
Fax: 562-279-2199
Email: shinhsvision.postech.ac.kr
Seq primer: T3 primer.
FEATURES
source
1. 142
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="F0-2D"
/lab_host="XL1-blue MRF"
/clone_lib="22 week old human fetal liver cDNA library"
/notes="Vector: pBluescriptII SK(-); Site 1: EcoRI; Site_2: XhoI; The cDNA library made by oligo-dT primed and
directionally cloned between 5'Exor I-XhoI3', sites."
BASE COUNT 27 a 39 c 38 g 38 t
ORIGIN
Alignment Scores: Length: 142
Pred. No.: 6.06e+03

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

```

RESULT 14
AW793767
LOCUS
DEFINITION
MRI-UM0009-290200-002-d07 UM0009 Homo sapiens cDNA, mRNA sequence.
AW793767
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 141)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=MRI-UM0009-290
200-002-d07&t3=2000-02-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 141.
FEATURES
Location/Qualifiers

```

Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 14 Gaps: 0

US-09-492-764B-20 (1-6) x R28953 (1-142)

QY 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 24 CCGTCTTAAACAAAA 7

Search completed: August 9, 2003, 19:11:34
 Job time : 1550 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: August 9, 2003, 18:20:24 ; Search time 47 Seconds
(without alignments)
56.347 Million cell updates/sec

Title: US-09-492-764B-20
Perfect score: 27
Sequence: 1 PXLKTK 6

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2_1/USPRO_pool/US0942764/runat_04082003_142327_11154/app_query.fasta_1.199
-DB=issued Patents_NA -QMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0942764@cgn1_1_44 @runat_04082003_142327_11154 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 12 rows of search results.

Table with columns: C, 13, 26, 96.3, 614, 4, US-09-702-705-1307, Sequence 1307, Ap. Contains 45 rows of alignment data.

ALIGNMENTS

RESULT 1
US-09-397-787-61
; Sequence 61, Application US/09397787
; Patent No. 6468750
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-61

Table with columns: Alignment Scores, Pred. No., Score, Percent Similarity, Best Local Similarity, Query Match, DB, Length, Matches, Conservative, Mismatches, Indels, Gaps. Contains summary statistics for the alignment.

```

;
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-153
;
; Alignment Scores:
; Pred. No.: 107 Length: 332
; Score: 26.00 Matches: 5
; Percent Similarity: 83.33% Conservative: 0
; Best Local Similarity: 83.33% Mismatches: 1
; Query Match: 96.30% Indels: 0
; DB: 4 Gaps: 0
;
; US-09-492-764B-20 (1-6) x US-09-062-451-153 (1-332)

```

```

;
; RESULT 2
; US-08-991-789A-153/c
; Sequence 153, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; REED, STEVEN G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 153:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 153:
;
; US-08-991-789A-153

```

```

;
; QY 1 Pro***LeuLysThrLys 6
; DB 247 CCGTCTTTAAAACTAAA 230
;
; RESULT 4
; US-09-598-326-153/c
; Sequence 153, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; REED, STEVEN G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
;
; US-09-492-764B-20 (1-6) x US-09-598-326-153 (1-332)

```

```

;
; QY 1 Pro***LeuLysThrLys 6
; DB 247 CCGTCTTTAAAACTAAA 230
;
; RESULT 3
; US-09-062-451-153/c
; Sequence 153, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; REED, STEVEN G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
;
; US-09-062-451-153 (1-332)

```

```

;
; Alignment Scores:
; Pred. No.: 107 Length: 332
; Score: 26.00 Matches: 5
; Percent Similarity: 83.33% Conservative: 0
; Best Local Similarity: 83.33% Mismatches: 1
; Query Match: 96.30% Indels: 0
; DB: 3 Gaps: 0
;
; US-09-492-764B-20 (1-6) x US-08-991-789A-153 (1-332)

```

```

;
; QY 1 Pro***LeuLysThrLys 6
; DB 247 CCGTCTTTAAAACTAAA 230
;
; RESULT 3
; US-09-062-451-153/c
; Sequence 153, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; REED, STEVEN G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
;
; US-09-062-451-153 (1-332)

```


TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 153:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 332 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 153:
 US-09-598-326-153

Alignment Scores:
 Pred. No.: 107 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservativity: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-598-326-153 (1-332)

Qy 1 Pro***LeuLysThrLys 6
 Db 247 CCGTCTTAAAAAACAATAA 230

RESULT 5

US-09-289-198-153/c
 Sequence 153 Application US/09289198
 Patent No. 6586570

GENERAL INFORMATION:
 APPLICANT: Frudakis, Tony N.
 APPLICANT: Smith, John M.
 APPLICANT: Reed, Steven G.
 APPLICANT: Misher, Lynda
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
 FILE REFERENCE: 210121.419C5
 CURRENT APPLICATION NUMBER: US/09/289,198
 CURRENT FILING DATE: 1999-04-09
 EARLIER APPLICATION NUMBER: US 09/062,451
 EARLIER FILING DATE: 1998-04-17
 EARLIER APPLICATION NUMBER: US 08/991,789
 EARLIER FILING DATE: 1997-12-11
 EARLIER APPLICATION NUMBER: US 08/838,762
 EARLIER FILING DATE: 1997-04-09
 EARLIER APPLICATION NUMBER: PCT/US97/00485
 EARLIER FILING DATE: 1997-01-10
 EARLIER APPLICATION NUMBER: US 08/700,014
 EARLIER FILING DATE: 1996-08-20
 EARLIER APPLICATION NUMBER: US 08/585,392
 EARLIER FILING DATE: 1996-01-01
 NUMBER OF SEQ ID NOS: 312
 SOFTWARE: FastSeq For Windows Version 3.0
 SEQ ID NO 153
 LENGTH: 332
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-289-198-153

Alignment Scores:
 Pred. No.: 107 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservativity: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-289-198-153 (1-332)

Qy 1 Pro***LeuLysThrLys 6
 Db 247 CCGTCTTAAAAAACAATAA 230

RESULT 6

US-08-276-452A-48/c
 Sequence 48 Application US/08276452A
 Patent No. 5646029
 GENERAL INFORMATION:
 APPLICANT: Chen, Chao-Guang
 APPLICANT: Mau, Shiao-Lim
 APPLICANT: Du, He
 APPLICANT: Gane, Alison M
 APPLICANT: Bacic, Antony
 APPLICANT: Clarke, Adrienne E
 TITLE OF INVENTION: Plant Arabinogalactan Protein (AGP) Genes
 NUMBER OF SEQUENCES: 91
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Greenlee and Winner, P.C.
 STREET: 5370 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: United States of America
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/276,452A
 FILING DATE: 18-JUL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Caruthers, Jennie M.
 REGISTRATION NUMBER: 34,464
 REFERENCE/DOCKET NUMBER: 27-91A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303)499-8080
 TELEFAX: (303)499-8089
 TELEX: 49617824
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 350 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..156
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..24
 OTHER INFORMATION: /note= "Derived amino acid sequence
 OTHER INFORMATION: matching the peptide sequence by protein
 OTHER INFORMATION: sequencing"
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1..20
 OTHER INFORMATION: /note= "Nucleotide sequence
 OTHER INFORMATION: corresponding to the Pca23P2a primers"
 US-08-276-452A-48

Alignment Scores:
 Pred. No.: 114 Length: 350
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservativity: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-08-276-452A-48 (1-350)

Qy 1 Pro***LeuLysThrLys 6
 Db 220 CCCACTCTCAAAAACAAA 203

RESULT 7
 US-08-798-744-48/c
 ; Sequence 48, Application US/08798744
 ; Patent No. 5830747
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Chao-Guang
 ; APPLICANT: Du, He
 ; APPLICANT: Mau, Shao-Lim
 ; APPLICANT: Gané, Alison M
 ; APPLICANT: Bacic, Antony
 ; APPLICANT: Clarke, Adrienne E
 ; TITLE OF INVENTION: Plant Arabinogalactan Protein (AGP) Genes
 ; NUMBER OF SEQUENCES: 91
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Greenlee and Winner, P.C.
 ; STREET: 5370 Manhattan Circle, Suite 201
 ; CITY: Boulder
 ; STATE: Colorado
 ; COUNTRY: United States of America
 ; ZIP: 80303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/798,744
 ; FILING DATE: 13-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/276,452
 ; FILING DATE: 18-JUL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Caruthers, Jennie M.
 ; REGISTRATION NUMBER: 34,464
 ; REFERENCE/DOCKET NUMBER: 27-91A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (303)499-8080
 ; TELEFAX: (303)499-8089
 ; TELEX: 49617824
 ; INFORMATION FOR SEQ ID NO: 48:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 350 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..156
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..24
 ; OTHER INFORMATION: /note= "Derived amino acid sequence
 ; OTHER INFORMATION: matching the peptide sequence by protein
 ; OTHER INFORMATION: sequencing"
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1..20
 ; OTHER INFORMATION: /note= "Nucleotide sequence
 ; OTHER INFORMATION: corresponding to the Pca23F2a primers"
 ; US-08-798-744-48
 Alignment Scores: Length: 350
 Pred. No.: 114 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 Indels: 0
 Gaps: 0

RESULT 8
 US-09-397-787-97/c
 ; Sequence 97, Application US/09397787
 ; Patent No. 6468758
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
 ; FILE REFERENCE: 210121.466C2
 ; CURRENT APPLICATION NUMBER: US/09/397,787
 ; CURRENT FILING DATE: 1999-09-16
 ; NUMBER OF SEQ ID NOS: 334
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 97
 ; LENGTH: 352
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-397-787-97
 Alignment Scores: Length: 352
 Pred. No.: 114 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 Indels: 0
 Gaps: 0

RESULT 9
 US-09-404-879A-113/c
 ; Sequence 113, Application US/09404879A
 ; Patent No. 6468546
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Algate, Paul A.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.462C2
 ; CURRENT APPLICATION NUMBER: US/09/404,879A
 ; CURRENT FILING DATE: 1999-09-24
 ; NUMBER OF SEQ ID NOS: 393
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 113
 ; LENGTH: 568
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-09-404-879A-113
 Alignment Scores: Length: 568
 Pred. No.: 197 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 Indels: 0
 Gaps: 0

RESULT 6
 US-09-492-764B-20 (1-6) x US-09-397-787-97 (1-352)
 OY 1 Pro***LeuLysthrlYs 6
 Db 219 CCATCTTTAAAACCTAAA 202
 Alignment Scores: Length: 352
 Pred. No.: 114 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 Indels: 0
 Gaps: 0

RESULT 6
 US-09-492-764B-20 (1-6) x US-09-404-879A-113 (1-568)
 OY 1 Pro***LeuLysthrlYs 6
 Db 219 CCATCTTTAAAACCTAAA 202
 Alignment Scores: Length: 568
 Pred. No.: 197 Matches: 5
 Score: 26.00
 Percent Similarity: 83.33%
 Best Local Similarity: 83.33%
 Query Match: 96.30%
 Indels: 0
 Gaps: 0

```

Db 455 CCACTTTAAAAACTAAA 438
RESULT 10
US-09-938-933-113/c
; Sequence 113, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; EARLIER FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-113
Alignment Scores:
Pred. No.: 197 Length: 568
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservatative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
US-09-492-764B-20 (1-6) x US-09-338-933-113 (1-568)
Qy 1 Pro***LeuLyThrLys 6
||| |||||
Db 455 CCACTTTAAAAACTAAA 438
RESULT 11
US-09-215-681-113/c
; Sequence 113, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-113
Alignment Scores:
Pred. No.: 197 Length: 568
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservatative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 4 Gaps: 0
US-09-492-764B-20 (1-6) x US-09-215-681-113 (1-568)
Qy 1 Pro***LeuLyThrLys 6
||| |||||
Db 455 CCACTTTAAAAACTAAA 438
RESULT 12
US-09-385-982-278

```

```

; Sequence 278, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDESE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 278
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(595)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-278
Alignment Scores:
Pred. No.: 208 Length: 595
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservatative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 3 Gaps: 0
US-09-492-764B-20 (1-6) x US-09-385-982-278 (1-595)
Qy 1 Pro***LeuLyThrLys 6
||| |||||
Db 167 CCACTTTAAAAACTAAA 184
RESULT 13
US-09-702-705-1307/c
; Sequence 1307, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liguin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1307
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(614)
; OTHER INFORMATION: n = A,T,C or G
US-09-702-705-1307
Alignment Scores:

```

```

Pred. No.:      215      Length:      614
Score:          26.00    Matches:      5
Conservative:   83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches:  1
Query Match:    96.30%  Indels:      0
DB:             4       Gaps:        0

```

US-09-492-764B-20 (1-6) x US-09-702-705-1307 (1-614)

```

Oy      1 Pro***LeuLysThrLys 6
      ||| ||| ||| ||| ||| |||
Db      389 CCTACACTCAAAACAAAA 372

```

RESULT 14

```

US-09-736-457-1307/c
; Sequence 1307, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, AiJun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1307
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(614)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-1307

```

```

Alignment Scores:
Pred. No.:      215      Length:      614
Score:          26.00    Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches:  1
Query Match:    96.30%  Indels:      0
DB:             4       Gaps:        0

```

US-09-492-764B-20 (1-6) x US-09-736-457-1307 (1-614)

```

Oy      1 Pro***LeuLysThrLys 6
      ||| ||| ||| ||| ||| |||
Db      389 CCTACACTCAAAACAAAA 372

```

RESULT 15

```

US-08-401-908-17/c
; Sequence 17, Application US/08401908
; Patent No. 5684146
; GENERAL INFORMATION:
; APPLICANT: Yoshinobu OKUNO et al.
; TITLE OF INVENTION: A DNA WHICH CODES FOR THE VARIABLE REGION OF
; TITLE OF INVENTION: ANTIBODY
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.

```

```

; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/401.908
; FILING DATE: March 10, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-401-908-17

```

```

Alignment Scores:
Pred. No.:      222      Length:      630
Score:          26.00    Matches:      5
Percent Similarity: 83.33%  Conservative: 0
Best Local Similarity: 83.33%  Mismatches:  1
Query Match:    96.30%  Indels:      0
DB:             1       Gaps:        0

```

US-09-492-764B-20 (1-6) x US-08-401-908-17 (1-630)

```

Oy      1 Pro***LeuLysThrLys 6
      ||| ||| ||| ||| ||| |||
Db      494 CCAGCGCTAAAGACAAAA 477

```

Search completed: August 9, 2003, 19:23:42
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 9, 2003, 18:20:24 ; Search time 118 Seconds
(without alignments)
104.898 Million cell updates/sec

Title: US-09-492-764B-20

Perfect score: 27

Sequence: 1 PXLTKK 6

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09492764/runat_04082003_142328_11184/app_query.fasta_1.199
-DB=Published Applications NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09492764@CGN_11_164_@runat_04082003_142328_11184
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
c 1	26	96.3	162 9	US-09-864-761-28130 Sequence 28130, A

c 2	26	96.3	173 10	US-09-783-590-1438
c 3	26	96.3	243 9	US-09-876-889-61
c 4	26	96.3	272 10	US-09-796-692-3839
c 5	26	96.3	272 14	US-10-040-862-3839
c 6	26	96.3	288 13	US-10-040-739-16
c 7	26	96.3	323 14	US-10-198-846-10216
c 8	26	96.3	332 9	US-09-810-936-153
c 9	26	96.3	332 10	US-09-429-755-153
c 10	26	96.3	332 10	US-09-924-400-153
c 11	26	96.3	332 10	US-09-796-692-6087
c 12	26	96.3	332 14	US-10-040-862-6087
c 13	26	96.3	332 14	US-10-212-679-153
c 14	26	96.3	339 14	US-10-198-846-12449
c 15	26	96.3	351 10	US-09-960-352-4606
c 16	26	96.3	352 9	US-09-876-889-97
c 17	26	96.3	362 11	US-09-918-995-19011
c 18	26	96.3	379 10	US-09-960-332-828
c 19	26	96.3	381 10	US-09-960-352-5632
c 20	26	96.3	405 10	US-09-960-352-1988
c 21	26	96.3	410 9	US-09-922-217-680
c 22	26	96.3	410 10	US-09-833-263-680
c 23	26	96.3	410 13	US-10-025-380-680
c 24	26	96.3	413 11	US-09-918-995-17229
c 25	26	96.3	416 10	US-09-954-456-1761
c 26	26	96.3	420 13	US-10-027-632-183086
c 27	26	96.3	425 10	US-09-960-352-9111
c 28	26	96.3	428 10	US-09-998-598-1121
c 29	26	96.3	434 10	US-09-960-332-6506
c 30	26	96.3	434 10	US-09-998-598-985
c 31	26	96.3	449 13	US-10-027-632-286623
c 32	26	96.3	449 13	US-10-027-632-286624
c 33	26	96.3	450 10	US-09-884-767A-216
c 34	26	96.3	453 9	US-09-764-869-2152
c 35	26	96.3	453 14	US-10-091-504-2152
c 36	26	96.3	458 13	US-10-027-632-75731
c 37	26	96.3	458 13	US-10-027-632-75732
c 38	26	96.3	458 13	US-10-027-632-76360
c 39	26	96.3	458 13	US-10-027-632-76361
c 40	26	96.3	458 13	US-10-027-632-290067
c 41	26	96.3	459 11	US-09-918-995-31847
c 42	26	96.3	459 13	US-10-027-632-56732
c 43	26	96.3	459 13	US-10-027-632-290068
c 44	26	96.3	459 13	US-10-027-632-290069
c 45	26	96.3	460 13	US-10-027-632-278022

ALIGNMENTS

RESULT 1

US-09-864-761-28130/c
 ; Sequence 28130, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 28130
 ; LENGTH: 162
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC004381.1
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.84
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
 ; OTHER INFORMATION: SWISSPROT HIT: P45676, EVALUE 2.30e+00
 ; OTHER INFORMATION: NT HIT: AFI26145.1, EVALUE 1.00e-31
 ; US-09-864-761-28130

Alignment Scores:
 Pred. No.: 148 Length: 162
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 9 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-864-761-28130 (1-162)

QY 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 63 CCCTCTCTGAAACCAAG 46

RESULT 2
 ; US-09-783-590-1438/c
 ; Sequence 1438, Application US/09783590
 ; Patent No. US20020110850A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Patrick J.
 ; APPLICANT: Haseltine, William A.
 ; APPLICANT: Li, Haodong
 ; APPLICANT: Rosen, Craig A.
 ; APPLICANT: Ruben, Steven M.
 ; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
 ; FILE REFERENCE: PO-16.2C1
 ; CURRENT APPLICATION NUMBER: US/09/783,590
 ; CURRENT FILING DATE: 2000-02-15
 ; PRIOR APPLICATION NUMBER: 08/420,856
 ; PRIOR FILING DATE: 1995-04-12

; PRIOR APPLICATION NUMBER: 08/346,731
 ; PRIOR FILING DATE: 1994-11-21
 ; NUMBER OF SEQ ID NOS: 12485
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1438
 ; LENGTH: 173
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (3)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (8)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (22)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (42)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (80)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (82)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (116)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (123)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (137)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (140)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc feature
 ; LOCATION: (156)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-783-590-1438

Alignment Scores:
 Pred. No.: 159 Length: 173
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 10 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-783-590-1438 (1-173)

QY 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 61 CCGCCCTTAAACCAAG 44

RESULT 3
 ; US-09-876-889-61
 ; Sequence 61, Application US/09876889
 ; Patent No. US20020076715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
 ; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
 ; FILE REFERENCE: 210121.466C3
 ; CURRENT APPLICATION NUMBER: US/09/876,889
 ; CURRENT FILING DATE: 2001-06-06
 ; NUMBER OF SEQ ID NOS: 353

SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 61
 LENGTH: 243
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-876-889-61

Alignment Scores:
 Pred. No.: 232 Length: 243
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 9 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-876-889-61 (1-243)

Oy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 145 CCAGCCTTAAGACCAAG 162

RESULT 4
 US-09-796-692-3839/c
 Sequence 3839, Application US/09796692
 Publication No. US20020198362A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Algate, Paul A.
 APPLICANT: Mannion, Jane
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY

FILE REFERENCE: 2077.001200
 CURRENT APPLICATION NUMBER: US/09/796,692
 CURRENT FILING DATE: 2001-03-01
 PRIOR APPLICATION NUMBER: 60/186,126
 PRIOR FILING DATE: 2000-03-01
 PRIOR APPLICATION NUMBER: 60/190,479
 PRIOR FILING DATE: 2000-03-17
 PRIOR APPLICATION NUMBER: 60/200,545
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: 60/200,303
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 60/200,779
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 60/206,201
 PRIOR FILING DATE: 2000-05-01
 PRIOR APPLICATION NUMBER: 60/200,999
 PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084
 PRIOR FILING DATE: 2000-05-04
 PRIOR APPLICATION NUMBER: 60/206,201
 PRIOR FILING DATE: 2000-05-22
 PRIOR APPLICATION NUMBER: 60/218,950
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: 60/222,903
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: 60/206,201
 PRIOR FILING DATE: 2000-08-04
 PRIOR APPLICATION NUMBER: 60/223,416
 PRIOR FILING DATE: 2000-08-04
 PRIOR APPLICATION NUMBER: 60/223,378
 PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9597
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 3839
 LENGTH: 272
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-796-692-3839

Alignment Scores:
 Pred. No.: 263 Length: 272
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 10 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-876-889-61 (1-243)

Oy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 145 CCAGCCTTAAGACCAAG 162

RESULT 5
 US-10-040-862-3839/c
 Sequence 3839, Application US/10040862
 Publication No. US20030078396A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Mannion, Jane
 APPLICANT: Retter, Marc
 APPLICANT: Corixa Corporation
 TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther.

US-09-492-764B-20 (1-6) x US-09-796-692-3839 (1-272)
 Oy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 52 CCCTCCCTCAAAACTAAA 35

RESULT 5
 US-10-040-862-3839/c
 Sequence 3839, Application US/10040862
 Publication No. US20030078396A1
 GENERAL INFORMATION:
 APPLICANT: Gaiger, Alexander
 APPLICANT: Mannion, Jane
 APPLICANT: Retter, Marc
 APPLICANT: Corixa Corporation
 TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther.

FILE REFERENCE: 014058-013520US
 CURRENT APPLICATION NUMBER: US/10/040,862
 CURRENT FILING DATE: 2001-11-06
 PRIOR APPLICATION NUMBER: US 60/186,126
 PRIOR FILING DATE: 2000-03-01
 PRIOR APPLICATION NUMBER: US 60/190,479
 PRIOR FILING DATE: 2000-03-17
 PRIOR APPLICATION NUMBER: US 60/200,545
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 60/200,303
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: US 60/200,779
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: US 60/200,999
 PRIOR FILING DATE: 2000-05-01
 PRIOR APPLICATION NUMBER: US 60/202,084
 PRIOR FILING DATE: 2000-05-04
 PRIOR APPLICATION NUMBER: US 60/206,201
 PRIOR FILING DATE: 2000-05-22
 PRIOR APPLICATION NUMBER: US 60/218,950
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 60/222,903
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: US 60/223,416
 PRIOR FILING DATE: 2000-08-04
 PRIOR APPLICATION NUMBER: US 60/223,378
 PRIOR FILING DATE: 2000-08-07
 PRIOR APPLICATION NUMBER: US 09/796,692
 NUMBER OF SEQ ID NOS: 10467
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 3839
 LENGTH: 272
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-040-862-3839

Alignment Scores:
 Pred. No.: 263 Length: 272
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 DB: 14 Gaps: 0

US-09-492-764B-20 (1-6) x US-10-040-862-3839 (1-272)

Oy 1 Pro***LeuLysThrLys 6
 ||| |||||
 Db 52 CCCTCCCTCAAAACTAAA 35

RESULT 6
 US-10-040-739-16
 Sequence 16, Application US/10040739

```

; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; LaValle, Edward
; Racie, Lisa
; Merber, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-040-739-16
Alignment Scores: Length: 288
Pred. No.: 280 Matches: 5
Score: 26.00 Conservat: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 13

US-09-492-764B-20 (1-6) x US-10-040-739-16 (1-288)
RESULT 7
QY 1 Pro***LeuLyThrLys 6
DB 130 CCTTCACTGAAACAAA 147
US-10-198-846-10216/c
; Sequence 10216, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049

```

```

; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10216
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-10216
Alignment Scores: Length: 323
Pred. No.: 317 Matches: 5
Score: 26.00 Conservat: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 14

US-09-492-764B-20 (1-6) x US-10-198-846-10216 (1-323)
QY 1 Pro***LeuLyThrLys 6
DB 208 CCTTCTTTAAAACAAA 191
RESULT 8
US-09-810-936-153/c
; Sequence 153, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Mishler, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 153
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-810-936-153
Alignment Scores: Length: 332
Pred. No.: 327 Matches: 5
Score: 26.00 Conservat: 0
Percent Similarity: 83.33% Mismatches: 1
Best Local Similarity: 83.33% Indels: 0
Query Match: 96.30% Gaps: 0
DB: 9

US-09-492-764B-20 (1-6) x US-09-810-936-153 (1-332)
QY 1 Pro***LeuLyThrLys 6
DB 247 CCGTCTTTAAAACAAA 230
RESULT 9
US-09-429-755-153/c
; Sequence 153, Application US/09429755A
; Patent No. US2002011467A1
; GENERAL INFORMATION:

```


; APPLICANT: Frudakis, Tony N.
 ; APPLICANT: Smith, John M.
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Misher, Lynda
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.419C6
 ; CURRENT APPLICATION NUMBER: US/09/429,755A
 ; CURRENT FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: -315
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 153
 ; LENGTH: 332
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-429-755-153

Alignment Scores:
 Pred. No.: 327 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservatative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-429-755-153 (1-332)

Qy 1 Pro***LeuLysThrLys 6
 Db 247 CCGTCTTTAAAAACTAAA 230

RESULT 10
 US-09-924-400-153/c
 ; Sequence 153, Application US/09924400
 ; Patent No. US20020165371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frudakis, Tony N.
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Smith, John M.
 ; APPLICANT: Misher, Lynda E.
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A. W.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Deng, Ta
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.419C12
 ; CURRENT APPLICATION NUMBER: US/09/924,400
 ; CURRENT FILING DATE: 2001-08-07
 ; NUMBER OF SEQ ID NOS: 340
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 153
 ; LENGTH: 332
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-924-400-153

Alignment Scores:
 Pred. No.: 327 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservatative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-924-400-153 (1-332)

Qy 1 Pro***LeuLysThrLys 6
 Db 247 CCGTCTTTAAAAACTAAA 230

RESULT 11
 US-09-796-692-6087
 ; Sequence 6087, Application US/09796692
 ; Publication No. US20020198362A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
 ; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
 ; FILE REFERENCE: 2077.001200
 ; CURRENT APPLICATION NUMBER: US/09/796,692
 ; CURRENT FILING DATE: 2001-03-01
 ; PRIOR APPLICATION NUMBER: 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 9597
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 6087
 ; LENGTH: 332
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (328)
 ; OTHER INFORMATION: n=A,T,C or G
 ; US-09-796-692-6087

Alignment Scores:
 Pred. No.: 327 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservatative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-796-692-6087 (1-332)

Qy 1 Pro***LeuLysThrLys 6
 Db 198 CCAACTTCAAAACCAAAA 215

RESULT 12
 US-10-040-862-6087
 ; Sequence 6087, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander

Alignment Scores:
 Pred. No.: 327 Length: 332
 Score: 26.00 Matches: 5
 Percent Similarity: 83.33% Conservatative: 0
 Best Local Similarity: 83.33% Mismatches: 1
 Query Match: 96.30% Indels: 0
 Gaps: 0

US-09-492-764B-20 (1-6) x US-09-796-692-6087 (1-332)

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.419C14
CURRENT FILING DATE: 2002-08-02
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 153
LENGTH: 332
ORGANISM: Homo sapiens
US-10-212-679-153

Alignment Scores:
Pred. No.: 327 Length: 332
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 14 Gaps: 0

US-09-492-764B-20 (1-6) x US-10-212-679-153 (1-332)

Oy 1 Pro***LeuLysThrLys 6
Db 247 CCGTCTTTAAAACTAAA 230

RESULT 14

US-10-198-846-12449/c
; Sequence 12449, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12449
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 334, 335, 336, 337, 338, 339
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12449

Alignment Scores:
Pred. No.: 335 Length: 339
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 14 Gaps: 0

US-09-492-764B-20 (1-6) x US-10-198-846-12449 (1-339)

Oy 1 Pro***LeuLysThrLys 6
Db 106 CCAAGCCTCAAAACCAAG 89

RESULT 15

US-09-960-352-4606/c
; Sequence 4606, Application US/09960352

APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy of Hematological Malignancies

FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 6087
LENGTH: 332
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (328)
OTHER INFORMATION: n=A,T,C or G
US-10-040-862-6087

Alignment Scores:
Pred. No.: 327 Length: 332
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 14 Gaps: 0

US-09-492-764B-20 (1-6) x US-10-040-862-6087 (1-332)

Oy 1 Pro***LeuLysThrLys 6
Db 198 CCAACTCTCAAAACCAAA 215

RESULT 13

US-10-212-679-153/c
; Sequence 153, Application US/10212679
; Publication No. US20030125536A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary
; APPLICANT: Hirst, Shannon Kathleen
; APPLICANT: Dillon, Davin
; APPLICANT: Foy, Teresa
; APPLICANT: Houghton, Ray
; APPLICANT: Persing, David
; APPLICANT: Kalos, Michael

```

; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4606
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3058-016-Q1-K1-E7
US-09-960-352-4606

```

```

Alignment Scores:
Pred No.: 348 Length: 351
Score: 26.00 Matches: 5
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 1
Query Match: 96.30% Indels: 0
DB: 10 Gaps: 0

```

US-09-492-764B-20 (1-6) x US-09-960-352-4606 (1-351)

```

Qy 1 Pro***LeuLysThrLys 6
Db 302 CCCTCCTTGAGACAAAA 285

```

Search completed: August 10, 2003, 06:17:27
Job time : 120 secs

THIS PAGE BLANK (USPTO)